Cross-Layer Design of FDD-OFDM Systems based on ACK/NAK Feedbacks

Zuleita K. M. Ho, *Member, IEEE,* Vincent K. N. Lau, *Senior Member, IEEE,* Roger S. K. Cheng *Member, IEEE*

arXiv:0905.4700v1 [cs.IT] 28 May 2009

Abstract— It is well-known that cross-layer scheduling which adapts power, rate and user allocation can achieve significant gain on system capacity. However, conventional cross-layer designs all require channel state information at the base station (CSIT) which is difficult to obtain in practice. In this paper, we focus on cross-layer resource optimization based on ACK/NAK feedback flows in OFDM systems without explicit CSIT. While the problem can be modeled as Markov Decision Process (MDP), brute force approach by policy iteration or value iteration cannot lead to any viable solution. Thus, we derive a simple closed-form solution for the MDP cross-layer problem, which is asymptotically optimal for sufficiently small target packet error rate (PER). The proposed solution also has low complexity and is suitable for realtime implementation. It is also shown to achieve significant performance gain compared with systems that do not utilize the ACK/NAK feedbacks for cross-layer designs or cross-layer systems that utilize very unreliable CSIT for adaptation with mismatch in CSIT error statistics. Asymptotic analysis is also provided to obtain useful design insights.

Index Terms— ACK, Acknowledgement, Cross-Layer, Feedback, Scheduling, Markov Decision Process, MDP, No CSI, Power Adaptation, Rate Adaptation

I. INTRODUCTION

A. Background and motivation

Cross-layer scheduling has been shown to achieve a significant performance gain in wireless systems as a result of multiuser diversity gain. Most of the existing cross-layer designs heavily rely on either perfect CSIT [6] [13][14] or imperfect [15] [18]/ delayed CSIT [7] [19].

1) Absence of Accurate CSIT and CSIT error statistics: Perfect CSIT is difficult to obtain in practice, especially in FDD systems in which explicit feedback is required. With imperfect CSIT [1], systematic packet errors would result even if powerful error correction codes are applied. This is because given the imperfect CSIT, there is uncertainty on the instantaneous mutual information at the base station and the scheduled data rate may exceed the instantaneous mutual information, leading to packet errors (channel outage) despite the use of powerful error correction coding. It has been shown [5][20] that packet errors cause significant degradation in cross-layer performance. There are some works to take into account of the imperfect CSIT or limited CSIT feedback in cross-layer design. For example, in [16] [17], the authors studied the

cross-layer design with noiseless limited feedback. In [15] [19], the authors studied OFDMA cross layer design with outdated CSIT. However, in all these works, the CSIT obtained is either noiseless (or no delay) or the statistics of the CSIT errors is assumed to be known [7]. However, in practice, the knowledge of CSIT errors statistics such as CSIT error variance and CSIT delay is needed and this is not easy to obtain because it depends on the mobility of the users as well as the multipath profile. It is quite challenging to have a robust cross-layer scheduling solution without the knowledge of CSIT error variance. On the other hand, regardless of the CSIT, there are always ACK/NAK flows between the mobiles (MS) and the basestations (BS). A robust cross-layer scheduling should make the best use of the ACK/NAK information which is embedded in the protocol. [2]

2) Accomodation of mobiles with different receiver capability: Conventional cross-layer design that utilized CSIT to perform resource allocation is essentially an open-loop system because BS cannot determine if the packet is received correctly or not even with the knowledge of CSIT (due to decoding errors). In practice, the system may have heterogeneous mix of mobiles with different capabilities (e.g. some has turbo decoding capability while some only has simple detection capability). To accommodate the heterogeneous mixture of receiver capability in the resource allocation, the BS has to rely on ACK/NAK flows (because the ACK/NAK flows give information about whether a packet can be decoded successfully or not). This closed loop information cannot be obtained in CSIT-based scheduler.

3) Heuristic Approach in existing literature: Recognizing the importance of utilizing the ACK/NAK in the resource allocation at the BS, there are existing works that discuss power control using ACK/NAK feedbacks. However, most of the works either considered power control on a wireless link only as well as utilizing heuristic algorithms or study the performance by simulation. For example, a power adaptation design and performance study utilizing ACK/NAK feedbacks for point-to-point systems have appeared in [21]-[24]. Cross-layer scheduling utilizing ACK/NAK feedbacks was investigated in [9] [10][25]. In particular, power control, rate adaptation and user scheduling for flat fading channels and frequency selective channels were carried out in [9] and [10] respectively whereas a rate adaptation scheme based on ACK/NAK feedbacks was proposed in [25]. The authors proposed a 2-level hierarchy stochastic scheduling algorithm based on *learning automata* (LA) for an AWGN channel by

---

[1]There are two meanings behind "imperfect CSIT" in the literature. The first meaning of imperfect CSIT refers to partial knowledge of CSIT such as limited feedback but the partial CSIT knowledge is received accurately (without errors) or timely (no delay). On the other hand, the second meaning of imperfect CSIT refers to inaccurate knowledge of CSIT (either with CSIT errors or outdatedness). In this paper, the term "imperfect CSIT" refers to the second meaning.

[2]in a similar way as *outerloop power control* in CDMA systems.

rate adaptation. Although the algorithm was shown to converge to the true channel state values, the convergence is not proven to maximize the throughput which is of usual practical concern. Moreover, in all these works [9][10][25], the algorithm designs are based on heuristic solutions and it is not clear what the best possible performance from the ACK/NAK information is. Furthermore, the suboptimal solutions obtained have high complexity and is not suitable for real-time implementations. Moreover, in all these existing designs, there is no mechanism to control the per-user packet error rate PER to a given target level. Yet, being able to control the PER of the wireless sessions per user is very important from the requirements of applications (e.g. voice and video codec).

Motivated by all the reasons above, we propose a robust closed-loop cross-layer design for OFDM systems where no explicit CSIT knowledge is needed at the base station. The cross-layer power allocation, user assignment as well as rate allocation are adaptive to the built-in 1-bit ACK/NAK feedbacks [1] [2] [3] from the selected users. Being built in at the link layer of most wireless systems and hence, the ACK/NAK feedbacks add no incremental cost to the proposed closed-loop design. Moreover, since the cross-layer solution is driven by the ACK/NAK feedbacks, it introduces robustness on the cross-layer performance with respect to uncertainty at the CSIT and propagation parameters. These robustness cannot be obtained by utilizing explicit limited CSIT feedback. However, there are several challenges in solving the problem:

B. Technical Challenges

1) Issues of packet errors: Conventional cross-layer optimization only consider sum ergodic capacity as the optimization objection. Ergodic capacity only considers the b/s/Hz transmitted by the BS regardless of packet errors. As a result, ergodic capacity is a reasonable performance metric only when the packet error is negligible (which is the case with perfect CSIT and very strong coding). However, in our case without CSIT, there is always systematic packet errors (due to channel outage) and this cannot be alleviated by just using strong coding. To accommodate packet errors, we have to use system goodput (b/s/Hz successfully received by the mobiles) as our performance metric. Note that goodput reduces to ergodic capacity in the case of no errors but in general, to deal with goodput, we need to deal with the cdf of mutual information (rather than the first order moment only) and this impose some technique challenges to the problem.

2) Issue of the MDP complexity: While the problem belongs to MDP, it is well-known that there is usually no simple solution (even numerically) using standard value-iteration and policy-iteration solutions (see details in section II). For instance, the MDP belongs to the class of infinite state space and brute-force approach has exponential complexity in the number of time slots M and hence, they could not give useful solutions. Instead of brute-force solution, we exploit some special structure of the OFDM and obtained a low complexity closed-form solution, which is asymptotically optimal for sufficiently small PER target.

3) Asymptotic Performance: As pointed out, all existing solutions are heuristic in nature and studied performance purely by simulations. This is because of the challenging nature of the problem. In this paper, we shall derive some asymptotic properties on the system performance so as to obtain some design insights.

C. Summary of Contributions

We consider the downlink of a wireless system with a base station and K mobile users over frequency selective fading channels (OFDM). The base station shall adapt the downlink rate, power and user selection in an OFDM system based on the ACK/NAK feedbacks from the mobiles. To take into account of potential packet errors due to channel outage, we consider an average system goodput which measures the number of bits successfully transmitted as our performance measure. The robust cross-layer design is modelled as a Markov Decision Process (MDP) [4] [35] [36] [37] with power, rate and user selection policies as the optimization variables so as to optimize the average system goodput while maintaining a target PER. It is well-known that MDP-based problems [26][27] always require complex value iteration algorithms. However, in this paper, we shall derive a simple closed-form solution for the MDP cross-layer problem which is asymptotically optimal for sufficiently small target PER. The proposed solution has low complexity and is suitable for realtime implementation. It is also shown to achieve significant performance gain compared with systems that do not utilize the ACK/NAK feedbacks for cross-layer designs or cross-layer systems that utilize very unreliable CSIT for adaptation with mismatch in CSIT error statistics. Furthermore, since the ACK/NAK feedbacks are generated by the mobiles based on CRC checking after packet detection, the proposed closed-loop cross-layer scheme is very flexible in the sense that it can automatically accommodate mobiles with different receive sensitivities in the RF or variations in the baseband estimation and decoding algorithms. Hence, the proposed scheme achieve significant goodput gain with built-in robustness against channel fluctuations as well as variations across the capabilities of different mobile receivers.

II. A Review on Markov Decision Process

MDP has found applications in ecology, economics and communications engineering since 1950 [28]. MDP is a modeling tool which describes a sequential decision making process. It is used to make the *optimal* sequence of decisions where outcomes of the problem are partly random and partly depend on such decisions. The advantage of MDP is that it provides a systematic framework for analysis of optimality, existence, dynamics and convergence of solutions.

A complete description of a MDP problem involves a *decision epoch*, a *state space*, a *control policy*, a *state transition kernel* as well as a *reward function*. The time line is first divided into decision epochs in which the controller makes decisions on control actions and the system receives *rewards* at the decision epochs. Specifically, at the m-th decision epoch, the system occupies a state $s_m \in \mathbb{S}$ where $\mathbb{S}$ denotes the state

space. Based on the observation on the causal state sequence $s_1, ..., s_m$, the controller takes a control action $\alpha_m \in \mathbb{A}$ where $\mathbb{A}$ is the set of actions. A *control policy* π is defined to be the *set of actions* for all possible state sequence. Based on the action α_m and the current state s_m, the system receives a reward $R(s_m, \alpha_m)$ and moves to the next state s_{m+1} according to the *state transition probability kernel* $P(s_m, \alpha_m, s_{m+1})$. The optimization problem is to find the optimal control policy so as to maximize the total rewards: $\arg\max_\pi \sum_{m=1}^M R(s_m, \alpha_m)$. As a result, a MDP problem can be characterized by the tuple $(\mathbb{S}, \mathbb{A}, P(.,.,.)), R(.,.))$. One reason why the MDP problem is difficult is due to the huge dimensions of the variable, namely the entire policy space π. As a result, a key step in solving the MDP is known as *divide-and-conquer*. Specifically, instead of optimizing for the entire problem, it can shown that the MDP can be solved by optimization of actions α_m on a per-stage basis.

There has been a lot of in-depth analysis of MDP [28] [29] and different branches of the problem. Different analysis are needed for finite state space problems v.s. infinite state space problems; finite horizon problems v.s. infinite horizon problems; *unconstrained* MDP v.s. *constrained MDP* etc. By constrained MDP, we mean that the problem has one or more constraints on the feasible policy space π. Constrained MDP problems are closely related to communication problems [29] such as power and rate control problems with an average delay constraint [30]; scheduling problems involving routing in ad-hoc networks [32] or handoff problems [31]. For example, in [31], the authors optimized the occurrence of path optimizations for inter-switch handoffs in wireless ATM networks. The expected total cost per call, including the switching/ handoff cost and signaling costs, is modeled as a infinite-horizon semi-Markov decision process [33] with discount rate. This expected total cost is the objective function to be minimized. At each decision epoch, the decision maker can choose to do path optimization or not which is modeled in the action set. Using divide-and-conquer principle, the MDP problem can be solved using *value iteration algorithm* or policy iteration algorithm [34]. The model is then extended to have QoS constraints.

This paper is outlined as follows. The channel model is firstly presented in section III. In section IV, the problem formulation is given as a cross-layer optimization problem and a MDP problem. The conventional solutions of MDP is provided at the end of section IV. The proposed solution, which is asymptotically optimal, is presented in section V. Simulation results are analyzed in the section VII. Section VIII presents the conclusion.

III. CHANNEL MODEL

We consider a downlink cross-layer scheduling problem in a frequency selective, block fading (in frequency) and quasi static (in time) channel. The bandwidth is divided into D frequency blocks. The fading gain in each frequency block is flat. With the use of OFDM, the fading of each frequency block is independent to other frequency blocks. Also, in the time domain, we assume that the channel remains quasi-static
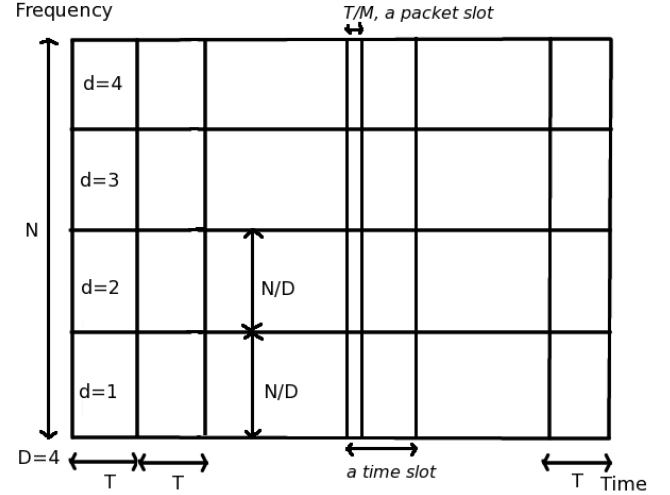

Fig. 1. The channel model is represented graphically. In the frequency domain, assume $D = 4$ frequency blocks within N subcarriers, there are $\frac{N}{D}$ subcarriers in each frequency block and have the same frequency gains. In the time domain, channel remains unchanged within T seconds: a time slot. M packets are transmitted in a time slot. Each packet consume $\frac{T}{M}$ seconds: a packet slot.

for a period of time T seconds and we call this a time slot. Thus, the fading gains on each frequency block remain the same throughout a *time slot*. Within a time slot, we send M packets which occupy the same amount of time, a *packet slot*, $\frac{T}{M}$ seconds. From now on, the names packet slot and slot are used interchangeably. With frequency block fading, there are N frequency sub-carriers in which $\lfloor \frac{N}{D} \rfloor$ frequency sub-carriers having the same fading gains form a block and there are D blocks in total. The fading gain represented by each frequency block is assumed to be independent of the other blocks. The model is summarized in figure 1.

Denote the number of users in the systems by K. Each user k sees a vector channel $\bar{h}_k = [h_{k,1}, \ldots, h_{k,D}]$ where $h_{k,j}$ is the channel power of frequency block j of user k. Stacking all vector channels, we have a channel power matrix H.

$$H = \begin{pmatrix} \bar{h}_1 \\ \bar{h}_2 \\ \vdots \\ \bar{h}_K \end{pmatrix} = \begin{pmatrix} h_{1,1} & h_{1,2} & \cdots & \cdots & h_{1,D} \\ h_{2,1} & h_{2,2} & \cdots & \cdots & h_{2,D} \\ \vdots & \ddots & & & \vdots \\ h_{K,1} & h_{K,2} & & \ddots & h_{K,D} \end{pmatrix}$$ (1)

Note that each entry $h_{k,d}$ is exponentially distributed with unit mean and variance. Denote the ACK/NAK feedback from each user k during packet slot m by $v_{k,m}$. Then,

$$v_{k,m} = \begin{cases} 1, & \text{ACK is received from user } k \text{ in slot } m; \\ 0, & \text{NAK is received from user } k \text{ in slot } m. \end{cases}$$ (2)

where ACK is received when the packet m is successfully decoded and NAK is received when the packet m has error.

The closed-loop cross-layer scheduler is as shown in figure 2. There are three optimization parameters, namely the user selection a_m, power level p_m and rate r_m. The parameters are determined for each packet m. At the receiver side, *each user k*
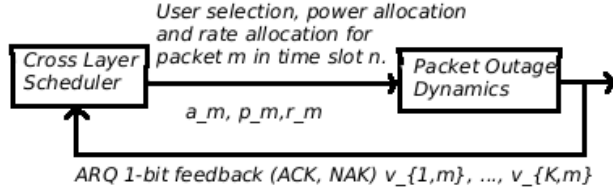
Fig. 2. Closed Loop Cross-Layer Scheduler. The user, power and rate optimization at the BSs is solely based on the 1-bit feedbacksfrom MSs.
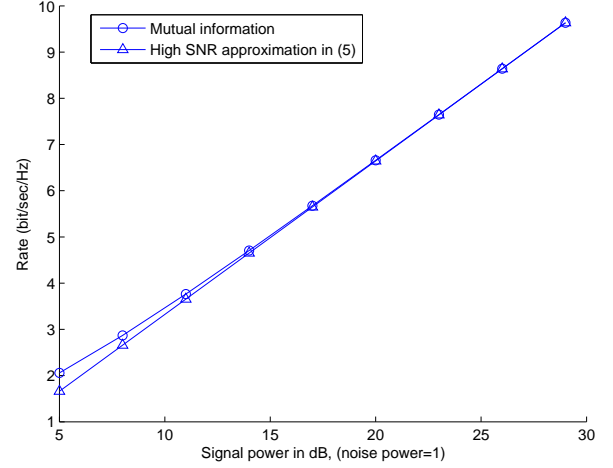

Fig. 3. Rate difference between mutual information and its approximation in (5). The difference is less than 2% in common operating region, between 10 to 30 dB.

would decode the packet and send a 1-bit ACK/NAK feedback $v_{k,m}$ to the transmitter. In m-th packet slot, the maximum achievable rate in bits is

$$c(p_m, \bar{h}_{a_m}) = \frac{NT}{DM} \sum_{d=1}^{D} \log_2(1 + \frac{p_m h_{a_m,d}}{N}) \qquad (3)$$

where noise power is normalized to be one.

Now, we can rewrite equation (2) mathematically,

$$v_{k,m} = \begin{cases} 1, & r_m \leq c(p_m, \bar{h}_{a_m}); \\ 0, & r_m \geq c(p_m, \bar{h}_{a_m}). \end{cases} \qquad (4)$$

In high SNR environment, the maximum bits per packet slot in equation (3) can be approximated by

$$c(p_m, \bar{h}_{a_m}) = \frac{NT}{DM} \sum_{d=1}^{D} \log_2(1 + \frac{p_m h_{a_m,d}}{N})$$

$$\xrightarrow[highSNR]{} \frac{NT}{DM} \left(\sum_{d=1}^{D} \left(\log_2 \left(\frac{p_m}{N} \right) \right) + \log_2(X_{a_m}) \right)$$

$$= c(p_m, X_{a_m}) \qquad (5)$$

where $X_k = \prod_{d=1}^{D} h_{k,d}$. This approximation significantly reduces the complexity of the system as the D-dimensional channel power gain vector is replaced by a scaler. In figure 3, we show the difference between the maximum bits per packet slot and its approximation in (5). The approximation error is less than 2% when the SNR is around 10dB.

Define the cumulative density function (CDF) of the random variable X_k to be

$$\phi(\chi) = Pr(X_k \leq \chi) \qquad (6)$$

which can be computed offline. Note that X_k is unknown to the transmitter which updates the set of all possible values of X_k in each packet slot m by the feedback $v_{k,m}$. The set of all possible values X_k, based on information received through feedbacks before packet slot m, is

$$\mathbb{X}_{k,m+1} = \begin{cases} \mathbb{X}_{k,m} \bigcap \{X_k : c(p_m, X_k) \geq r_m\}, & v_{k,m} = 1; \\ \mathbb{X}_{k,m} \bigcap \{X_k : c(p_m, X_k) < r_m\}, & v_{k,m} = 0. \end{cases} \qquad (7)$$

For example, at packet slot 1, $m = 1$, the set of real channel power gains for $\mathbb{X}_{k,1}$ is all real numbers $\mathbb{R}^+$. A pair of power and rate (p_1, r_1) is selected. A packet is broadcasted with power p_1 and rate r_1. At the end of packet slot 1, ACK/NAK feedbacks $v_{k,1}$ for all users k are received. $\mathbb{X}_{k,2}, \forall k$ are then updated using (7). At the end of packet slot 2, $\mathbb{X}_{k,3}$ are updated accordingly and so on. Note that the set $\mathbb{X}_{k,m}$, as described in (7), would solely depend on the causal power allocation, rate allocation and ACK/NAK feedbacks from the users.

IV. PROBLEM FORMULATION

This section is targeted to reveal the mathematical description of the optimization problem. The problem is best explained by first writing down the optimization variables which are the power, rate and user selection policies defined in the following. We would then provide the mathematical expression of the system goodput which is the optimization objective in this paper. A problem statement and its corresponding mathematical representation are provided. A subsection is given here to explain the transformation of the optimization problem to a MDP problem.

A. Problem formulation as a cross-layer optimization problem

For simplicity, denote the causal user assignments, rate sequence and power sequence from slots 1 to $m - 1$ by $A_m = (a_1, a_2, \ldots, a_{m-1})$, $R_m = (r_1, r_2, \ldots, r_{m-1})$ and $P_m = (p_1, p_2, \ldots, p_{m-1})$ respectively. Also, denote the causal ACK/NAK feedbacks for slots 1 to $m - 1$ from all users by

the matrix V_m

$$V_m = \begin{pmatrix} v_{1,1} & v_{1,2} & \cdots & \cdots & v_{1,m-1} \\ \vdots & \ddots & & & \vdots \\ v_{K,1} & v_{K,2} & & \ddots & v_{K,m-1} \end{pmatrix}$$

$$= \begin{pmatrix} \bar{v}_1^m \\ \vdots \\ \bar{v}_K^m \end{pmatrix} = (\bar{v}^1, \bar{v}^2, \ldots, \bar{v}^{m-1}) \quad (8)$$

Definition 1 (Power Allocation Policy): A power allocation policy

$$\mathcal{P} = \left\{ (p_m)_{V_m} : \sum_{m=1}^{M} p_m = P_0 \right\} \quad (9)$$

is defined as the set of all power allocation at the m-th packet slot where $m \in [1, M]$. The subscript notation $(.)_{V_m}$ denotes that the power allocation at the m-th packet slot is a function of the ACK/NAK feedbacks up to the $(m-1)$-th packet slot V_m. The power allocation policy $\mathcal{P}$ is restricted by the total power constraint P_0.

Similarly, we define the rate allocation policy and user selection policy.

Definition 2 (Rate Allocation Policy): A rate allocation policy

$$\mathcal{R} = \left\{ (r_m)_{V_m} : r_m \in \mathbb{R}^+ \right\} \quad (10)$$

is defined as the set of all rate allocation at m-th packet slot where $m \in [1, M]$ and $\mathbb{R}^+$ is the set of all positive real numbers. The policy is determined by causal ACK/NAK feedbacks up to slots $m - 1$.

Definition 3 (User Selection Policy): A user selection policy

$$\mathcal{A} = \{ (a_m)_{V_m} : a_m \in \{1, \ldots, K\} \} \quad (11)$$

is defined as the set of all user selection at m-th packet slot where $m \in [1, M]$. The policy is determined by the causal ACK/NAK feedback sequences up to slots $m - 1$. The user selection at m-th packet slot a_m denotes the index of user selected.

Let the feedback of user a_m at packet slot m in time slot z be $v_{a_m,m}(z)$. The number of packet errors in time slot z equals to the sum of packet errors of the M packets sent within time slot z: $\sum_{m=1}^{M}(1 - v_{a_m,m}(z))$. The total number of packet errors in Z time slots is $\sum_{z=1}^{Z} \sum_{m=1}^{M}(1 - v_{a_m,m}(z))$. Thus, the packet error rate averaged over time slots is

$$P_e = \lim_{Z \to \infty} \frac{1}{MZ} \sum_{z=1}^{Z} \sum_{m=1}^{M}(1 - v_{a_m,m}(z)). \quad (12)$$

As the channel gain remains quasi-static within a time slot and is independent of that in other time slot, the averaged packet error rate can be written as the expectation of number of packet errors within a time slot over channel realizations.(We drop the notation of time slot z)

$$P_e = \mathbf{E}_H \frac{1}{M} \sum_{m=1}^{M}(1 - v_{a_m,m}) \quad (13)$$

where $\mathbf{E}_H(.)$ denotes expectation over the random variable H.

Note that the packet error rate can be simplified as follows.

$$P_e = Pr(c(p_m, X_{a_m}) < r_m) \quad (14)$$

The average system goodput $\bar{G}$ (averaged over ergodic samples of time slots) is given by:

$$\bar{G}(\mathcal{P}, \mathcal{R}, \mathcal{A}) = \mathbf{E}_H \left\{ \sum_{m=1}^{M} v_{a_m,m} r_m \right\}$$

$$= \sum_{m=1}^{M} Pr(c(p_m, X_{a_m}) > r_m) r_m. \quad (15)$$

In most wireless systems, a target packet error rate (PER) is assigned due to various application requirements. Let ϵ be that PER. For example, the PER, ϵ, is of the order of 10^{-2} for voice applications. The relation between $\mathbb{X}_{k,m}$ and ϵ (5) is given by

$$1 - \epsilon = Pr(c(p_m, X_{a_m}) \geq r_m | \mathbb{X}_{a_m,m})$$

$$= Pr(X_{a_m} \geq \theta_m | \mathbb{X}_{a_m,m}) \quad (16)$$

where

$$\theta_m = \left(\frac{N}{p_m} \right)^D 2^{\frac{r_m D M}{NT}}. \quad (17)$$

To conclude, the cross-layer optimization problem can be formulated as

Problem 1 (Cross-layer formulation): Determine the optimal power allocation policy $\mathcal{P}$, rate allocation policy $\mathcal{R}$ and user assignment $\mathcal{A}$ so as to maximize the average system goodput $\bar{G}(\mathcal{P}, \mathcal{R}, \mathcal{A})$ subject to the target PER requirement $1 - \epsilon = Pr(X_{a_m} \geq \theta_m | \mathbb{X}_{a_m,m})$ and the total power constraint $\sum_{m=1}^{M} p_m \leq P_0$.

The optimization problem above is difficult to solve due to the huge dimension of variables involved. Yet, we shall illustrate below that the total system goodput $\bar{G}$ can be expressed recursively and hence, the problem above can be expressed as a Markov Decision Problem. Define $F_m(\bar{P}_m, W_{m-1})$ to be the maximized goodput sum from slot m to M (from packet slot m to the last packet slot) subject to power constraint $\bar{P}_m$ and *causal* power allocations, rate allocations and feedbacks from users i.e.

$$F_m(\bar{P}_m, W_{m-1}) = \max_{\mathbf{p}_m, \mathbf{r}_m, \mathbf{a}_m} \mathbf{E}_H \left\{ \sum_{i=m}^{M} v_{a_i,i} r_i \right\} \quad (18)$$

where $W_{m-1} = (V_{m-1}, A_{m-1}, \Theta_{m-1} = (\theta_1, \ldots, \theta_{m-1}))$ and $\mathbf{p}_m$ denotes the vector of power allocation from p_m to p_M. Similar notations apply to $\mathbf{r}_m$ and $\mathbf{a}_m$. The maximization is subject to the PER requirement $Pr(X_{a_m} \geq \theta_m | \mathbb{X}_{a_m,m}) = 1 - \epsilon$ and the total power constraint $\sum_{i=m}^{M} p_i \leq \bar{P}_m$. We first have the following lemma about $F_m(\bar{P}_m, W_{m-1})$.

Lemma 1: $F_m(\bar{P}_m, W_{m-1})$ can be espressed recursively as

$$F_m(\bar{P}_m, W_{m-1}) \quad (19)$$

$$= \max_{p_m, r_m, a_m} \left\{ (1 - \epsilon) r_m + \mathbf{E}_{V_m} \left[F_{m+1}(\bar{P}_m - p_m, W_m) \right] \right\}.$$

Proof: See subsection IX-A in appendix. ∎

Note that the maximization variables are p_m, r_m, a_m, the power, rate and user selection in packet slot m, instead of the selections from slot m till the last slot. As a result,

this facilitate the divide-and-conquer approach to the original optimization problem in (1).

From (15), the maximized system goodput is

$$\bar{G}^*(\mathcal{P}, \mathcal{R}, \mathcal{A}) = \max_{\mathbf{a}_m, \mathbf{p}_m, \mathbf{r}_m} \sum_{m=1}^{M} \mathbf{E}_H \{v_{a_m,m}\} r_m. \quad (20)$$

By definition of F_m in equation (18), the optimized goodput is

$$\bar{G}^*(\mathcal{P}, \mathcal{R}, \mathcal{A}) = F_1(P_0, W_0) \quad (21)$$

$$\text{subject to} \quad \sum_{m=1}^{M} p_m \leq P_0$$

$$Pr(X_{a_m} \geq \theta_m | X_{a_m} \in \mathbb{X}_{a_m,m}) = 1 - \epsilon$$

$$\theta_m = \left(\frac{N}{p_m}\right)^D 2^{\frac{r_m DM}{NT}}$$

where W_0 is a empty set. As a result, the optimized system goopdut $\bar{G}^*(\mathcal{P}, \mathcal{R}, \mathcal{A}) = F_1(P_0, W_0)$ can be obtained recursively from equation (19). We shall eleborate the recursive solution in the following sections.

B. Problem Formulation as a Markov Decision Process

As explained in Section II, a MDP problem is characterized by the tuple $(\mathbb{T}, \mathbb{S}, \mathbb{A}, P(s, \alpha, s'), R(s, \alpha))$. In our case, the decision epochs of the base station $\mathbb{T} = \{1, 2, ..., M\}$ corresponds to the scheduling slots. In the following, we shall discuss the association of our cross-layer optimization problem with the MDP tuple, namely the state space $\mathbb{S}$, action space $\mathbb{A}$, state transition kernel as well as the per-stage reward function. Based on that, we shall formally recast the problem into an MDP.

- **State Space Association** With $\Theta_m = [\theta_1, \ldots, \theta_m]$, define $U(\Theta_m, \bar{v}_k^m)$ and $L(\Theta_m, \bar{v}_k^m)$ to be the upper bound and lower bound of CSI which is some information gathered by the ACK/NAK feedbacks $\bar{v}_k^m$ and θ_m in equation (17). The state space, $\mathbb{S}$, is a collection of the following vectors s.

$$s = (L(\Theta_m, \bar{v}_k^m), U(\Theta_m, \bar{v}_k^m), \theta_m, \bar{P}_m, \bar{R}_m, \vec{s}^{(ACK)}, \vec{s}^{(NAK)}) \quad (22)$$

  where $\bar{P}_m$ is the remaining power; $\bar{R}_m$ is the sumrate from slot m to M, $\vec{s}^{(ACK)}$ and $\vec{s}^{(NAK)}$ are the *pointers* to the states if ACK:$v_m = 1$ and NAK: $v_m = 0$ respectively. The CSI can take all possible real values and therefore make the state space $\mathbb{S}$ infinite. However, as illustrated in an example in the following subsection, the decision tree built by state transitions in our problem is a lot smaller in size.

- **Action Space and Policy Association** The action taken at each state s consists of the selection of power p_m, transmission rate, r_m, and the user selection, a_m. The set of possible actions $\mathbb{A}$ at every state s is independent of decision epoch m and it is given by:

$$\mathbb{A} = \mathbb{A}_{s,m} = \{(p_m, r_m, a_m) \in \quad (23)$$
$$\{p \in \mathbb{R}^+ : p \leq P_0\} \times \mathbb{R}^+ \times \{1, \ldots, K\}\}.$$

- **State Transition Kernel Association** The transition probability $P(s, \alpha, s')$ is a real value function which maps $\{\mathbb{S} \times \mathbb{A} \times \mathbb{S}\}$ to $[0, 1]$. In our case, the probability of going from state s to state s' by action $\alpha \in \mathbb{A}$ is time invariant.

In each decision epoch, m, a selection of actions, α_m, takes place, meaning that the base station selects the power p_m and the transmission rate r_m to user a_m. After every user k receives the packet, each of them would decode the packet header and transmit a 1-bit feedback to base station, $v_{k,m}$. This 1-bit feedback carries the information of ACK (1) or NAK (0). The transition probability captures the probability of such ACK (1) or NAK (0) and would take the system to a different state. For instance, the current state is denoted by s; the state after receiving ACK s^a; the state after receiving NAK s^n. The probability of receiving ACK is $\mathcal{P}_a$ and that of NAK is $1 - \mathcal{P}_a$. The action taken is α. We have

$$P(s, \alpha, s^a) = \mathcal{P}_a; \quad (24)$$
$$P(s, \alpha, s^n) = 1 - \mathcal{P}_a. \quad (25)$$

And

$$\sum_{s' \in \mathbb{S}} P(s, \alpha, s') = 1 \quad (26)$$

The state transition probability is described in equation (31) in which θ' is the third element in s' and θ is the third element in s. The upper and lower bound of CSI would be modified according to the ACK/NAK feedbacks received. After updating the bounds, the probability of ACK, which is equal to the probability of the event that the channel power X_k lies between the lower bound and state θ', has to equal $1 - \epsilon$, as dictated by the error constraint. Evaluate the probability, we have equation (32).

- **Per-stage Reward** To decide which actions in $\mathbb{A}$ should be carried out, we would need a decision rule d_m. The decision rule d_m is a history-dependent function. Define the history δ_m to be a vector of past states, actions and feedbacks.

$$\delta_m = [s_1, \alpha_1, \ldots, s_{m-1}, \alpha_{m-1}, s_m] \quad (27)$$

The recursive relation is therefore

$$\delta_m = [\delta_{m-1}, \alpha_{m-1}, s_m]. \quad (28)$$

Denote the set of all histories by Δ_m. Note that

$$\Delta_1 = \mathbb{S} \quad (29)$$
$$\Delta_2 = \mathbb{S} \times \mathbb{A} \times \mathbb{S}$$
$$\vdots$$
$$\Delta_m = \mathbb{S} \times \mathbb{A} \times \cdots \times \mathbb{S}$$
$$= \Delta_{m-1} \times \mathbb{A} \times \mathbb{S}$$

The history dependent rule d_m maps Δ_m to $\mathbb{A}$.

A control policy is a plan specified by a sequence of decision rules. A control policy π is

$$\pi = (d_1, d_2, \ldots, d_M), \quad d_i \in \Delta_i, i = 1, \ldots, M \quad (30)$$

$$P(s, \alpha, s') = P(\theta_{m+1} = \theta' | \theta_m = \theta, \alpha_m = \alpha) = \begin{cases} \epsilon, & Pr(X_k > \theta' | X_k > L(\Theta_m, \bar{v}_k^m), X_k < U(\Theta_m, \bar{v}_k^m)) = \epsilon \\ 1 - \epsilon, & Pr(X_k > \theta' | X_k > L(\Theta_m, \bar{v}_k^m), X_k < U(\Theta_m, \bar{v}_k^m)) = 1 - \epsilon \\ 0 & \text{otherwise.} \end{cases} \tag{31}$$

$$P(s, \alpha, s') = \begin{cases} \epsilon, & \phi(\theta_{m+1} = \theta') = (1 - \epsilon)\phi(U(\Theta_m, \bar{v}_k^m)) + \epsilon\phi(L(\Theta_m, \bar{v}_k^m)) \\ 1 - \epsilon, & \phi(\theta_{m+1} = \theta') = \epsilon\phi(U(\Theta_m, \bar{v}_k^m)) + (1 - \epsilon)\phi(L(\Theta_m, \bar{v}_k^m)) \\ 0 & \text{otherwise.} \end{cases} \tag{32}$$

$$P(s_m, \alpha, s_{m+1}) = \begin{cases} \epsilon, & \phi(\theta_{m+1}) = (1 - \epsilon)\phi(U(\Theta_m, \bar{v}_k^m)) + \epsilon\phi(L(\Theta_m, \bar{v}_k^m)) \\ 1 - \epsilon, & \phi(\theta_{m+1}) = \epsilon\phi(U(\Theta_m, \bar{v}_k^m)) + (1 - \epsilon)\phi(L(\Theta_m, \bar{v}_k^m)) \\ 0 & \text{otherwise.} \end{cases} \tag{33}$$

The per-stage reward function is

$$R(s_m, \alpha_m) = \begin{cases} P(s_m, \alpha, s_{m+1,a})r_m & if\ v_m = 1; \\ 0 & if\ v_m = 0; \end{cases} \tag{34}$$

where $s_{m+1,a}$ denotes the state at slot $m + 1$ if s_m is reached at slot m and action α_m is taken.

Problem 2 (The MDP formulation): The MDP problem is defined as a maximization problem of the reward function, in our case, the system goodput $F_1(P_0, W_0)$. Thus, the problem statement is, with slightly abuse of notation

$$\max_{\pi} \left\{ \sum_{m=1}^{M} R(s_m, \alpha_m) \right\} \tag{35}$$

such that $\forall m = 1, \dots, M, s_m, s_{m+1} \in \mathbb{S}, \alpha_m \in \mathbb{A}, r_m \in \mathbb{R}^+$ and equation (33) is satisfied.
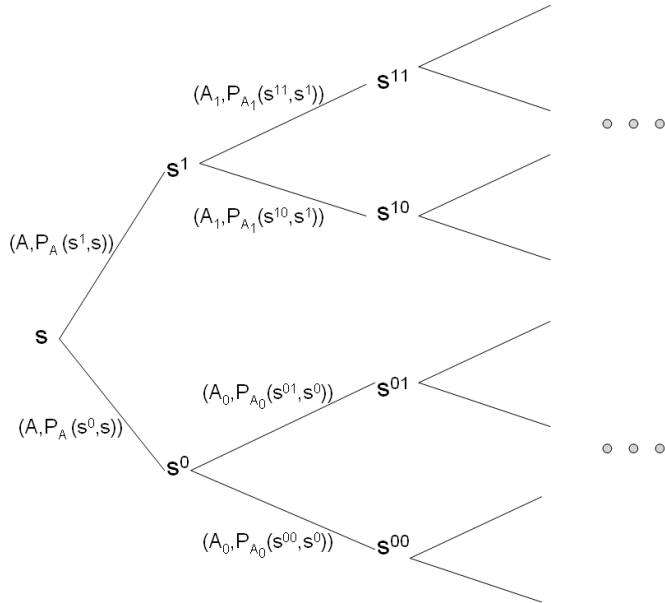

Fig. 4. A state transition diagram example. With only 2 possible outcomes at each state (node), the state space (the number of nodes) increases exponentially, hence the problem size.

C. A State Transition Example

To illustrate the state transition of a MDP, a state transition diagram assigned with an initial state is given in figure 4 by only drawing transition branches corresponding to the tuples of scheduled action and the corresponding non-zero transition probability. Note that this diagram only shows a fragment of the whole decision tree because there are more than one possible initial state.

The decision tree has $O(|\theta_m|)^3 \times 2^M$ elements, where $|\theta_m|$ is the number of values θ_m can take. In other words,

$$|\mathbb{S}| = O(|\theta_m|)^3). \tag{36}$$

There are $|\theta_m|$ possible values of the lower bound $L(\Theta_m, \bar{v}_k^m)$. For example, $L(\Theta_m, \bar{v}_k^m) \in \{y_1, \dots, y_{|\theta_m|}\}$ where $y_b < y_{b+1}$. For each value of lower bound y_b, there are $|\theta_m| - b - 1$ values of $U(\Theta_m, \bar{v}_k^m)$ and θ_m. Thus, the total number of possible states is $1^2 + 2^2 + \dots + |\theta_m|^2 = O(|\theta_m|^3)$.

With either positive or negative feedbacks, each state can only branch to 2 possible next states. Assume that we start on one of these states. The number of possible *descendents* would be equal to the sum of the series $1 + 2 + 2^2 + 2^3 + \dots + 2^{M-1}$ which is 2^M. Thus, the total number of nodes in the tree is $O(|\theta_m|)^3 \times 2^M$.

Denote the elements in the state space $\mathbb{S}$ by

$$\mathbb{S} = \left\{ s, \{s^0, s^1\}, \{s^{00}, s^{01}, s^{10}, s^{11}\}, \dots, \{s^{q_{M-1}}\} \right\} \tag{37}$$

where q_{M-1} denotes any possible binary sequence of length $M - 1$. The binary sequence represents the causal ACK or NAK feedbacks received. For example, state s^{00} represents that 2 NAKs have been received and state s^{101} represents that the first and the third transmission are correct and the second transmission or guess is incorrect. The state s^{q_i} is at the i-th level of the tree which means the $(i+1)$-th packet transmission (with the root being the zeroth level). In the diagram, only transitions with non-zero probability are drawn. The transition probability corresponding to action $A_{q_i} \in \mathbb{A}$ from state s^{q_i} to state $s^{[q_i,0]}$, meaning that a NAK is received at $(i+1)$-th packet transmission, is denoted by the probability $P\left(s^{q_i}, A_{q_i}, s^{[q_i,0]}\right)$. At each state s^{q_i}, there are two possible transition branches

$$ACK: \tag{38}$$
$$\left(A_{q_i}, P\left(s^{q_i}, A_{q_i}, s^{[q_i,1]}\right)\right) = ((p_{i+1}, r_{i+1}, a_{i+1}), 1 - \epsilon)$$
$$NAK: \tag{39}$$
$$\left(A_{q_i}, P\left(s^{q_i}, A_{q_i}, s^{[q_i,0]}\right)\right) = ((p_{i+1}, r_{i+1}, a_{i+1}), \epsilon).$$

D. Conventional Solutions of MDP

A conventional solution to a MDP consists of backward and forward recursions. The backward recursions set up a huge searching tree/ table which would involves dynamic programming. In the forward recursions, the system states evolve through the tree. Here we adopted the Finite Horizon-Policy Evaluation Algorithm in [28] for the backward recursion.

Algorithm 1 Conventional Finite Horizon- Policy Evaluation Algorithm

1: Each node in the tree consists of following fields: $(L, U, \theta_m, \bar{P}_m, \bar{R}_m, \vec{s}^{(ACK)}, \vec{s}^{(NAK)})$.
2: Initialization: $m \leftarrow M, \forall L, U, p_M, \theta_M$
$$F_M^*(p_M, \delta_M) = \max_{d_M(\delta_M)} Pr(c(p_M, \theta_M) > r_M)r_M$$
3: if $m = 1$, stop. Otherwise, go to step 4.
4: $m \leftarrow m - 1, \forall s_m, p_m, \bar{P}_m, L, U$
Evaluate $F_m^*(\bar{P}_m, \delta_m) = \max_{d_m(\delta_m)} \{P(s_m, \alpha, s_{m+1})r_m$
$+P(s_m, \alpha, s_{m+1})F_{m+1}^*(\bar{P}_m - p_m, \delta_m | v_{a_m,m} = 1)$
$+(1 - P(s_m, \alpha, s_{m+1}))F_{m+1}^*(\bar{P}_m - p_m, \delta_m | v_{a_m,m} = 0)\}$
such that the constraints in equation (33) are satisfied and $P(s_m, \alpha, s_{m+1}) = 1 - \epsilon$
5: (p_m, a_m, r_m) are given by $d_m(\delta_m)$ obtained in step 4.
6: $\bar{R}_m = F_m^*(\bar{P}_m, \delta_m)$ which is the accumulated rate of this node and its descendents.
7: $\vec{s}^{ACK}, \vec{s}^{NAK}$ are computed in (33)

After building up a table in backward recursion using algorithm 1, from $m = M \rightarrow 1$, we established a large binary tree with each node represents a particular estimate of channel power and each branch corresponds to an ACK/NAK feedback. Each path from the root to the leaves corresponds to a sequence of estimates and the corresponding feedbacks. In *Online Evolution* (algorithm 2), we read this tree from the root and traverse down to the leaves. Each packet is transmitted with parameters marked in that node and a new node is reached according to the ACK/NAK feedbacks.

Note that the drawback of such algorithm is that the requirement of memory is huge as there are numerous possible states. In our problem, the state space is infinite. Even if we discretize the state space as an approximation, the complexity of the brute-force approach has exponential complexity in M and hence, could not give viable solutions.

V. PROPOSED SOLUTIONS

The MDP can be solved by a backward recursion followed by a forward recursion. In this section, we shall first elaborate the backward recursive solution, namely the *Optimal State Evolution* followed by the forward recursion, namely the

---

Algorithm 2 Conventional Online State Evolution Algorithm

1: Set $m = 1$ and start state
$s = (0, \infty, \theta_m, P_0, \bar{R}_m, \vec{s}^{(ACK)}, \vec{s}^{(NAK)})$
where $\bar{R}_m$ is the maximum among the nodes with $L = 0, U = \infty$.
2: If $m = M + 1$, stop, otherwise go to step 3.
3: Transmit packets as prescribed by decision rule $d_m(\delta_m)$ computed in algorithm 1.
4: Receive an ACK/NAK feedback $v_{k,m}$ from each user k.
5: Update the upper and lower bound of CSI.
$L = \theta_m$ if $v_{a_m,m} = 1$
$U = \theta_m$ if $v_{a_m,m} = 0$
6: Evolve to next state according to the bounds of CSI $\vec{s}^{(ACK)}, \vec{s}^{(NAK)}$ and feedbacks $v_{k,m} \forall k$.
7: $m + 1 \leftarrow m$, go to step 2.

Online Envolution. Unlike conventional solution for MDP, we proposed a simple closed-form solution which is asymptotically optimal for sufficiently small PER. The proposed solution only has complexity $O(M)$, which is in big contrast with brute-force complexity $O(exp(M))$.

A. Optimal State Evolution

We illustrate how to combine the target PER ϵ, with the knowledge obtained from feedbacks to generate estimates of channel power θ_m. Note that θ_m in equation (17) is always either $\sup \mathbb{X}_{k,m}$ or $\inf \mathbb{X}_{k,m}$ as equation (7) can be rewritten as

$$\mathbb{X}_{k,m+1} = \begin{cases} \mathbb{X}_{k,m} \bigcap \{X_k : X_k \geq \theta_m\}, & v_{k,m} = 1; \\ \mathbb{X}_{k,m} \bigcap \{X_k : X_k < \theta_m\}, & v_{k,m} = 0. \end{cases} \tag{40}$$

The lower bound and upper bound of $\mathbb{X}_{k,m+1}$ are

$$L(\Theta_m, \bar{v}_k^m) = \max\{\theta_i : v_{k,i} = 1, 1 \leq i \leq m\} \tag{41}$$
$$U(\Theta_m, \bar{v}_k^m) = \min\{\theta_i : v_{k,i} = 0, 1 \leq i \leq m\}. \tag{42}$$

Combine (16) with the knowledge obtained from feedbacks:

$$Pr\left(X_k \geq \theta_{m+1} | X_k \geq L(\Theta_m, \bar{v}_k^m), X_k < U(\Theta_m, \bar{v}_k^m)\right) = 1 - \epsilon \tag{43}$$

Rearranging the terms in equation (43), we have the dynamics of θ_m

Lemma 2: At each packet slot m, the estimate of channel power X_{a_m} is computed by the causal feedbacks $\bar{v}_{a_m}^{m-1}$ and the lower and uppwer bound of X_{a_m}

$$\phi(\theta_m) = \epsilon\phi(U(\Theta_{m-1}, \bar{v}_{a_m}^{m-1})) + (1 - \epsilon)\phi(L(\Theta_{m-1}, \bar{v}_{a_m}^{m-1})) \tag{44}$$

where $\phi(\theta_m)$ is the cdf of X_{a_m} (6).
Proof: see section IX-B in appendix. ∎

B. User Selection

Evaluate the expectation in $F_m(\bar{P}_m, W_{m-1})$ defined in (19), we obtain equation (45). Solving equation (45), a stochastic programming tree would be needed. Yet, as ϵ is small for practice, the decision tree is reduced to equation (46).

The complexity of the problem is reduced from exponential to linearity with m.

$$F_m(\bar{P}_m, W_{m-1}) = \max_{p_m, r_m, a_m} \left\{(1-\epsilon)r_m + (1-\epsilon)F_{m+1}(\bar{P}_m - p_m, W_m|v_{a_m,m} = 1) + \epsilon F_{m+1}(\bar{P}_m - p_m, W_m|v_{a_m,m} = 0)\right\}$$
$$(45)$$

$$F_m(\bar{P}_m, W_{m-1}) = \max_{p_m, r_m, a_m} \left\{(1-\epsilon)r_m + (1-\epsilon)F_{m+1}(\bar{P}_m - p_m, W_m|v_{a_m,m} = 1)\right\}. \qquad (46)$$

$$d_m(\delta_m) = \left(p_m = \frac{\epsilon \bar{P}_m}{1 - (1-\epsilon)^{M-m+1}}, r_m = \frac{NT}{DM} \log_2\left(\left(\frac{p_m}{N}\right)^D \theta_m\right), a_m = \arg\max_k L(\Theta_{m-1}, \bar{v}_k^{m-1})\right) \qquad (47)$$

Lemma 3: The optimal user selection strategy

$$a_m = \arg\max_k L(\Theta_{m-1}, \bar{v}_k^{m-1}) \qquad (48)$$

of (46) is

Proof: See subsection IX-C in appendix. ∎

C. Power Allocation

Lemma 4: The power allocation policy

$$p_m = \frac{\epsilon \bar{P}_m}{1 - (1-\epsilon)^{M-m+1}} \qquad (49)$$

, where $\bar{P}_m = P_0 - \sum_{i=1}^{m-1} p_i$ is the remaining power at time m, is an optimal policy with respect to optimization problem (46).

Proof: See subsection IX-D in appendix. ∎

D. Rate Allocation

Given the causal feedback, power and rate information W_m and the channel estimate/state values θ_m in (59) at each slot m, the rate allocation is computed by the following

$$r_m = \frac{NT}{DM} \log_2\left(\left(\frac{p_m}{N}\right)^D \theta_m\right) \qquad (50)$$

E. Online Evolution

With new information, $v_{k,m-1}$ arrives in each slot m, we proceed on the decision tree according to the updated upper and lower bounds of CSI and the feedbacks. The set $\mathbb{X}_{k,m}$ is modified to contain only the possible values of the channel power gain based on the causal ACK/NAK feedbacks. $\mathbb{X}_{k,m} = \{x : L(\Theta_m, \bar{v}_k^m) < x < U(\Theta_m, \bar{v}_k^m)\}$ The transmission parameters according to the decision rule are in equation (47).

User a_m is selected such that she contains the largest possible channel power gain. As proved before, the power allocation is static and solely depends on the total power and the target error probability constraint. The data rate is adapted according to channel estimate θ_m and feedbacks $v_{a_m,m-1}$. The online scheduling policy is illustrated in figure 5.

VI. ASYMPTOTIC ANALYSIS

This section is devoted to prove that the goodput achieved in a packet slot would be equal to the instantaneous mutual information of the slot as if they were perfect CSIT when the number of transmissions or number of packet transmissions
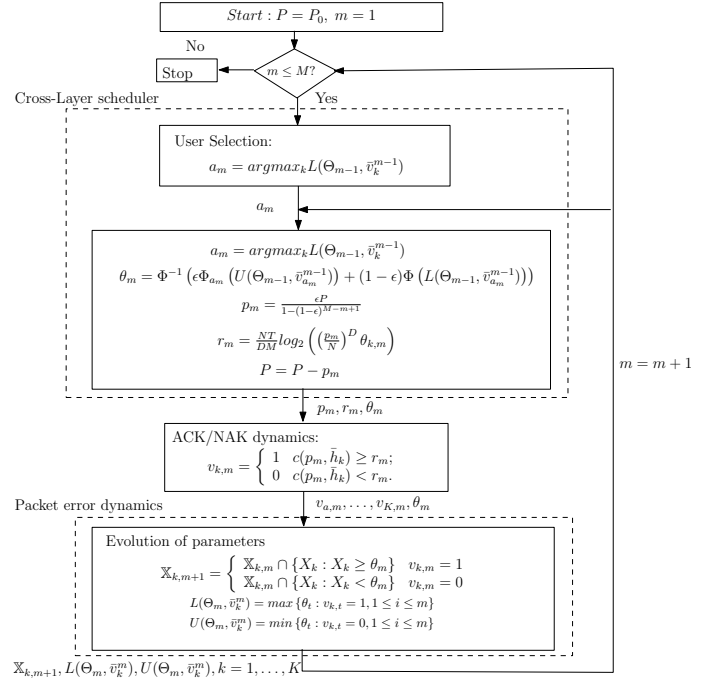

Fig. 5. Structure and implementation of the proposed solution.

tends to infinity. In other words, there is zero steady-state-error in the recursive solution. To prove such claim, we would need the following four theorems.

Lemma 5: At packet slot m, the users selection set $\mathbb{K}_m$ denotes the set of users who have the largest potential channel power gains.

$$\mathbb{K}_m = \{k : L(\Theta_m, \bar{v}_k^m) > L(\Theta_m \bar{v}_{k'}^m), \forall k' \notin \mathbb{K}_m\} \qquad (51)$$

The users selection set $\mathbb{K}_m$ at slot m is a subset of $\mathbb{K}_{m-1}$.

$$\mathbb{K}_m \subset \mathbb{K}_{m-1} \qquad (52)$$

The number of elements in $\mathbb{K}_m$ is $|\mathbb{K}_m|$ which decreases with m.

Proof: See subsection IX-E in appendix. ∎

Lemma 6: For all users k in user selection set $\mathbb{K}_m$ at each slot m, the channel power gains X_k have lower bounds and upper bounds $L(\Theta_m, \bar{v}_k^m)$ and $U(\Theta_m, \bar{v}_k^m)$.

Proof: See subsection IX-F in appendix. ∎

Lemma 7: Define the gap between the upper and lower bounds of channel power gains to be $w_m = U(\Theta_m, \bar{v}_{a_m}^m) - L(\Theta_m, \bar{v}_{a_m}^m)$. w_m monotonically decreases with m.
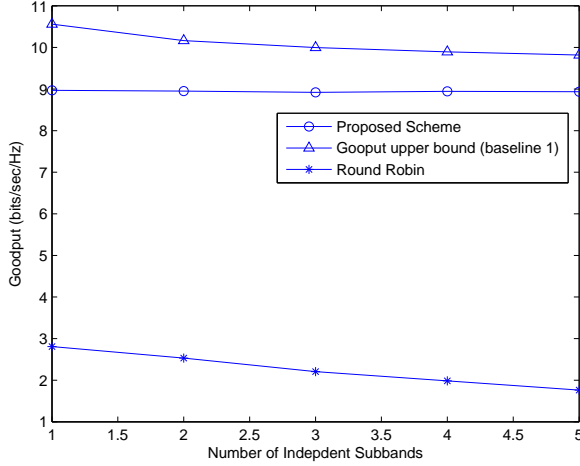
Proof: See subsection IX-G in appendix. ∎

Fig. 6. Average system goodput vs number of independent subbands with transmit SNR=30dB, $P_0 = 24W, K = 3, M = 30, PER = 0.05$.

Fig. 7. Average system goodput vs Average SNR with $P_0 = 24W, K = 3, D = 3, M = 30, PER = 0.05$. The proposed solution has the same slope as the upper bound (with perfect CSIT).

Lemma 8: When number of transmissions goes to infinity, the scheduled rate r_m achieves capacity of the system in perfect CSIT case. In the other words, the scheduled rate r_m is equal to the capacity achieved by selecting user which gives highest capacity and using perfect CSIT. Or mathematically,

$$\lim_{m \to \infty} r_m = \lim_{m \to \infty} \frac{NT}{DM} log_2 \left(\left(\frac{p_m}{N} \right)^D \theta_m \right) = c(p_m, X_{a_m}).$$

Proof: See subsection IX-H in appendix. ∎

VII. RESULTS AND DISCUSSIONS

In this section, we would discuss the simulation results with the following simulation settings. The bandwidth of the systems is 20 MHz which is divided into 64 subcarriers (N=64). Throughout these subcarriers, there are D group of independent subbands. The time slot $T = 0.1$ sec and we compared our proposed solution with two baselines. Specifically, in baseline 1, we assume the BS has perfect CSIT and performs standard power adaptation and hence, it serves as a goodput upper bound. In baseline 2, we consider round robin scheduling which does not utilize any CSIT information and hence, has very robust performance against CSIT errors. Note that the performance of baseline 1 is obtained under perfect CSIT assumption and therefore is not achievable. By comparing with baseline 1, we can guage how *optimal* the proposed solution could achieve. Similarly, by comparing with baseline 2 (which is a common approach in the absence of CSIT), we could guage the potential performance advantage that can be captured by utilizing the built-in ACK/NAK feedback flows.

A. Effects of Number of Independent Subbands

In figure 6, the sum of goodput in 30 packets transmitted is plotted against the number of independent subbands D with $P_0 = 24W, SNR = 30dB, K = 3$ and target $PER = 0.05$. Note that our proposed solution achieved 85% and 91% of the performance upper bound (baseline 1) when $D = 1$ and
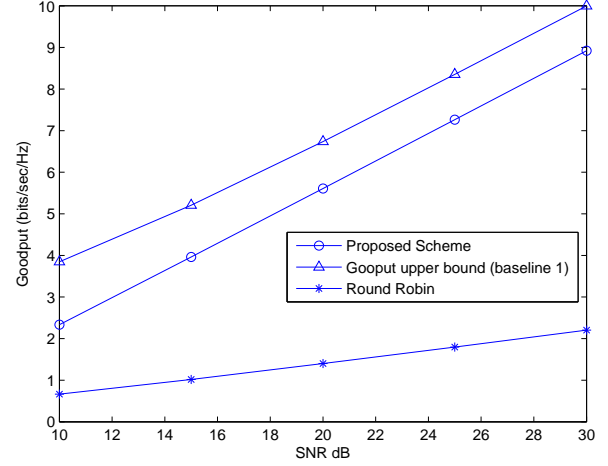
5 respectively. Compared with baseline 2 (RR), the proposed solution achieved very significant 500 % goodput gain. This illustrated the importance of utilizing the 1-bit ACK/NAK flows in the resource allocation.

Note that the goodput upper bound (baseline 1) decreases with D in figure 6 because the system did not take advantage of the frequency diversity as the selected user has to transmit on every frequency channels. When the number of independent channels increases, the capacity function, being concave in channel gains, decreases.

B. Effects of Transmit SNR

In figure 7, there are 3 users and each user has 3 independent channels. With transmission of 30 data packets in a time slot, the system goodput of the proposed solution achieves 60% and 89% of the performance upper bound (baseline 1) in low and high SNR scenarios respectively. Compared with baseline 2 (RR), the proposed solution has significant 400% gain in high SNR regime.

C. Effects of Number of Users

Figure 8 illustrates the system goodput vs number of users for $D = 3, M = 30, SNR = 30dB, P_0 = 24W$. Similarly, the proposed scheme achieved 93 % and 85 % of the performance upperbound (baseline 1) with 1 user and 9 users respectively. Compared with baseline 2 (RR), the proposed scheme achieved 400% goodput gain.

D. Effects of Target PER ϵ

Figure 9 illustrates the system goodput vs target PER for $SNR = 30dB, P_0 = 24W, K = 3, M = 30$ and $D = 3$. We observe that when the target PER is low, the proposed solution will be more conservative in determining the transmit data rate in order to avoid packet errors due to channel outage, On the other hand, when the target PER is high, the proposed
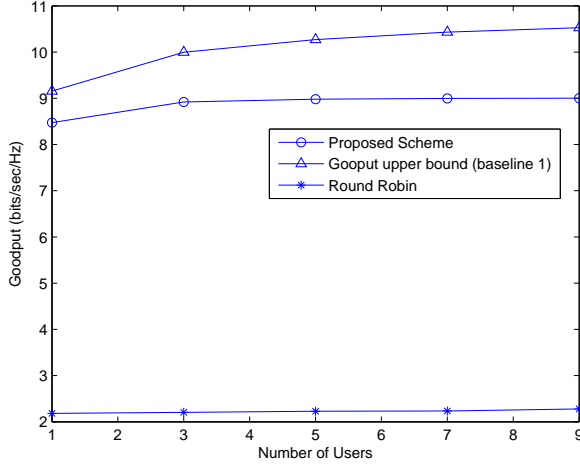
Fig. 8. Average system goodput vs numer of users with transmit SNR=30dB, $P = 24W, D = 3, M = 30, PER = 0.05.$: Capacity increases with number of users because of multi-user diversity, so as the proposed solution.
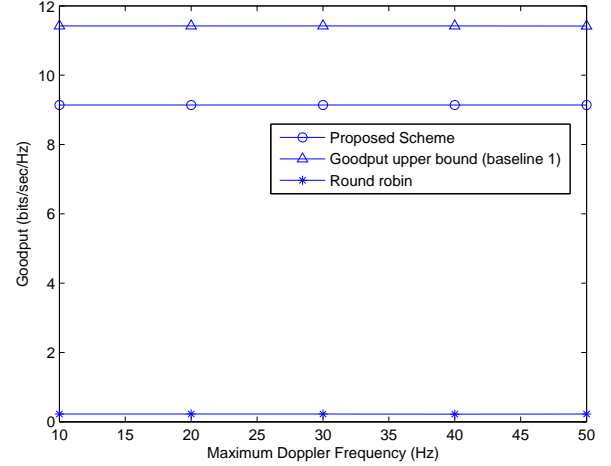

Fig. 10. Average system goodput vs maximum Doppler frequency. The users have i.i.d. random speed (uniformly distributed from 0 to $f_{d,max}$ throughout the simulation. $P_0 = 24W, K = 4, D = 3, M = 30, SNR = 30dB, PER = 0.1$
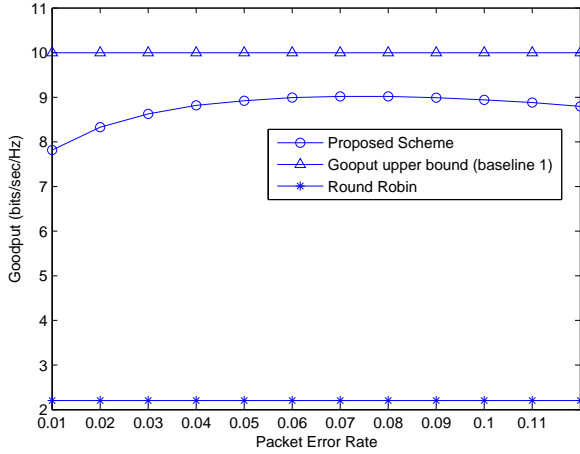

Fig. 9. Average system goodput v.s. target PER with transmit SNR=30dB, $K = 3, D = 3, M = 30$: With small target PER (e.g. errors sensitive applications), the proposed solution is conservative and acheive a less throughput. With high PER, the proposed solution may be over-optimistic on channel quality. In medium PER, the proposed solution gives the best performance.
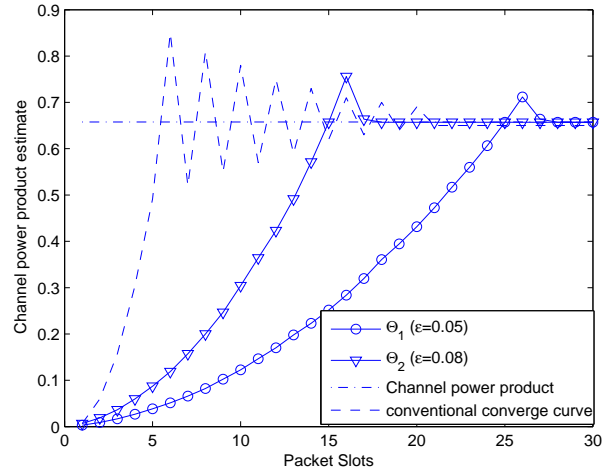

Fig. 11. Value of channel gains estimate Θ with Different PER targets in Different Packet Slots: The proposed solution maximizes goodput and therefore avoids over-estimating (resulting an NAK), hence the non-oscillating curve. A less target PER ϵ, which is more conservative, may prolong the convergence speed.

solution becomes more aggressive in transmitting data but the goodput will be limited by high channel outage probability. As a result, there is an optimal target PER, if one is interested to optimize the system goodput. Note that the performance upper bound of baseline 1 and the baseline 2 goodput performance is insensitive to the target PER.

E. Effects of Mobility

To study the robustness of the proposed scheme w.r.t. mobility, we assume the users have i.i.d. random speed (with Doppler frequency uniformly distributed from 0 to $f_{d,max}$). Figure 10 illustrates the average system goodput vs $f_{d,max}$ with $SNR = 30dB, P_0 = 24W, K = 4$ and $D = 3$. Observe that the proposed solution is quite robust even up to moderate

mobility of 50 Hz, which corresponds to 22.5 km/hr at 2GHz frequency. This robustness is due to the closed-loop feedback mechanism in the proposed solution.

F. Dynamics of Strategies

1) Tradeoff between Convergence Speed and Target PER: An example of the procedure of the algorithm is given in figures 11 to 14.

Figure 11 plots the channel power gain estimate θ_m in a particular channel realization v.s. time epoch m. ACK's are received until $m < 25$ and $m < 16$ for the curves PER $\epsilon = 0.5$ and 0.8 respectively. The upper bound of the θ_{a_m} is updated with NAK and θ_{a_m} converges to the true channel power gain
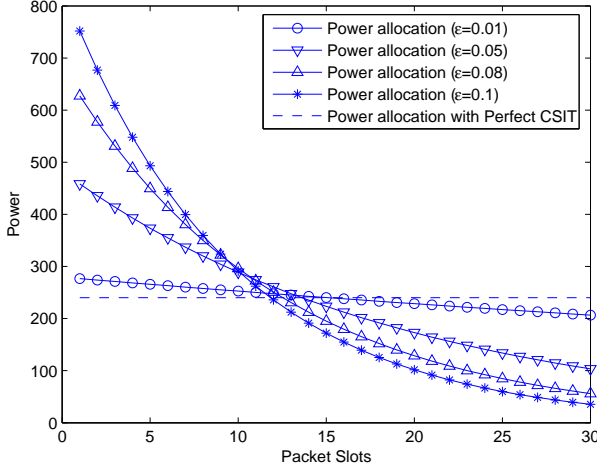
Fig. 12. Scheduled Power with Different PER Targets in Different Packet Slots

Fig. 13. Scheduled Data Rate with Different PER Targets in Different Packet Slots

product. The convergence time is shorter with high PER. It is because large PER provides larger flexibility for estimation. Yet, the throughput yield from large PER may be lower than that of small PER.

Moreover, conventional convergence curves would quickly climb close to the channel power gain product, overshoot, oscillate and then converge, as plotted in figure 11. The convergence curve of our scheduling scheme would not oscillate because any additional overshoot would waste power, time and the potential data transmission. Thus, our scheduling scheme increases steadily, overshoots once and converges.

2) Power Allocation Strategies for Different Outage Target: The power allocation of system with $P_0 = 24W, K = 3, D = 3, M = 30, SNR = 30dB$, is plotted in figure 12. Note that the power allocation strategies depend on the target PER ϵ. The objective is to maximize the goodput sum in all packet slots which can be separated into current goodput and future goodput as in equation (46). To maximize the goodput sum for large PER, more power should be allocated at the early slots to have as much successful transmission as possible . Notice that, as PER decreases, the power allocation converges to the power allocation for perfect CSIT, equal power allocation. It is because at the extreme case of zero PER, the probability of getting outage is zero, meaning that we have perfect CSIT (baseline 1).

3) Rate Allocation Strategies for Different PER Target: Assume $P_0 = 24W, SNR = 30dB, D = 3, K = 3, M = 30$. The rate allocation curves with different PER target are plotted in figure 13. Note that the area under the curve is the throughput. The data rate achieved by baseline 1 is plotted with a dotted line. Notice that the area achieved by small PER, 0.01, is small and the area increases by increasing the PER. However, area decreases after PER 0.07 which is the optimal PER in the current system assumption. An over-conservative PER target would yield too little goodput as the $\mathbb{X}_{a_m}$ is under estimated. An over-optimistic PER target would also yield a low goodput as outage occurs when $\mathbb{X}_{a_m}$ is over estimated.
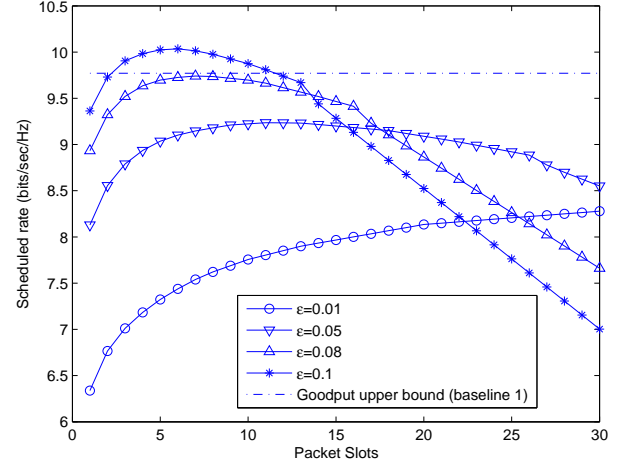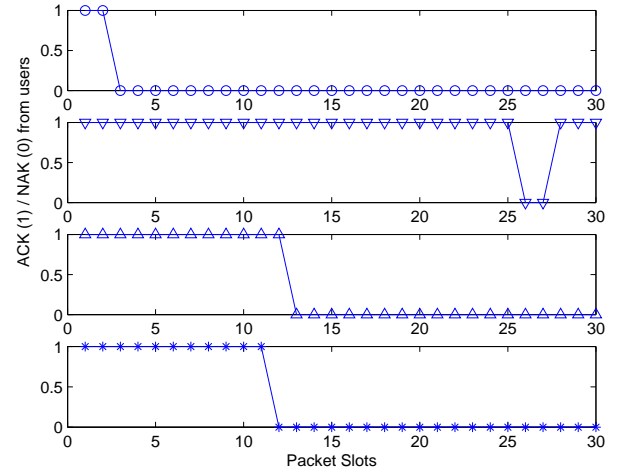

Fig. 14. Acknowledgements from Different Users, top (user 1), second one from the top (user 2) and so on

The allocated rate r_m increases with the increment of knowledge of the channel power gain in figure 13. Then r_m decreases after slot 10 because the scheduler has spent half of the total power in the first 10 slots. Less rate is resulted from smaller power remained for these 20 slots.

4) Acknowledgements Reveal CSIT: In figure 14, the acknowledgements from user 1 (from the top) to user 4 (from the bottom) are plotted whereas 1 denotes positive acknowledgement (ACK) and 0 denotes negative acknowledgement (NAK). After each transmission, each user decodes the packet header and feedback to transmitter. If a user k reports NAK at slot m, user k would have a channel power gain less than the channel power gain estimate at slot m, θ_m. Thus, we know that $\theta_2 \leq X_1 < \theta_3$, $\theta_{11} \leq X_4 < \theta_{12} \leq X_3 < \theta_{13}$. Since NAK are received at slot 25 and 26, we know that $\theta_{24} \leq X_2 < \theta_{27}$.

VIII. CONCLUSIONS

In this paper, we considered the OFDM resource optimization problem based on ACK/NAK feedbacks from the mobiles without explicit CSIT at the base station. We derive a simple closed-form solution for the MDP cross-layer problem which is asymptotically optimal for sufficiently small target PER. The proposed solution also has low complexity and is suitable for realtime implementation. Simulation results revealed that the system goodput performance of the proposed solution achieved 89% of the performance upper bound (perfect CSIT performance) and has over 400% gain compared to round robin scheduling. Due to the built in closed-loop feedback mechanism, the proposed scheme is shown to have robust performance against CSIT errors and different mobility. Asymptotic analysis is also provided to obtain useful design insights.

IX. APPENDIX

A. Recursive Property of Goodput

Recall from equation (19). Expectation over the channel power H is the same as the iterative expectation $E_{\hat{V}_m} E_{H|\hat{V}_m}$ where $\hat{V}_m$ is the feedbacks from users from slot m to M. Recall V_m, defined in (8) is the causal feedbacks from slot 1 to $m-1$. Combining V_m and $\hat{V}_m$ gives the whole history: $(V_m, \hat{V}_m) = V_M$.

$$F_m(\bar{P}_m, W_{m-1}) = \max_{\mathbf{p}_m, \mathbf{r}_m, \mathbf{a}_m} E_{\hat{V}_m} E_{H|\hat{V}_m} \left\{ \sum_{i=m}^{M} v_{a_i, i} r_i \right\}. \tag{53}$$

Evaluating the expectation yields

$$F_m(\bar{P}_m, W_{m-1}) = \max_{\mathbf{p}_m, \mathbf{r}_m, \mathbf{a}_m} E_{\hat{V}_m} \left\{ \sum_{i=m}^{M} Pr(c(p_i, X_{a_i}) > r_i) r_i \right\} \tag{54}$$

Separate the instantaneous goodput at slot m from the goodput sum from slot $m+1$ to M. Take an iterative expectation and obtain equation (63).

Since the first term does not depend on V_m nor v_m, it simplifies to (64).

Note that the second term is the expectation of $F_{m+1}(\bar{P}_m - p_m, W_m)$ over $\mathbf{v}_m$ according to equation (54). Equation (65) can be obtained.

B. Dynamics of θ_m

Denote the event $X_k \geq L(\Theta_{m-1}, \bar{v}_k^{m-1})$ by $\mathcal{L}$ and $X_k < U(\Theta_{m-1}, \bar{v}_k^{m-1})$ by $\mathcal{U}$ respectively. Employ the theorem of conditional probability on equation (43).

$$\frac{Pr(X_k \geq \theta_m, \mathcal{L}, \mathcal{U})}{Pr(\mathcal{L}, \mathcal{U})} = 1 - \epsilon \tag{55}$$

Recall the cdf of X_k, ϕ, in (6), (55) can be rewritten as

$$\frac{\phi(U(\Theta_{m-1}, \bar{v}_k^{m-1})) - \phi(\theta_m)}{\phi(U(\Theta_{m-1}, \bar{v}_k^{m-1})) - \phi(L(\Theta_{m-1}, \bar{v}_k^{m-1}))} = 1 - \epsilon \tag{56}$$

Rearranging the terms and equation (43) can be obtained.

C. Optimal User Selection

This section is to prove that the user selection $a_m = \arg\max_k L(\Theta_{m-1}, \bar{v}_k^{m-1})$ maximizes $F_m(\bar{P}_m, W_{m-1})$ in (46). Substitute $\theta_m = \left(\frac{N}{p_m}\right)^D 2^{\frac{DMr_m}{NT}}$ to $F_m(\bar{P}_m, W_{m-1})$ and we obtain equation (66).

Further expand (66), we obtatin (67)

As we assume $v_m, \dots, v_M = 1$, we have $\theta_m = L(\Theta_m, \bar{v}_{a_{m+1}}^m)$ and therefore

$$\theta_{m+1} = \phi^{-1} \left(\epsilon \phi(U(\Theta_m, \bar{v}_{a_{m+1}}^m)) + (1 - \epsilon)\phi(\theta_m) \right). \tag{57}$$

As $\phi(\theta_m)$ is the CDF of θ_m, $\phi(\theta_m)$ is monotonic increasing with θ_m, so as ϕ^{-1}. Thus, θ_{m+1} increases with θ_m. According to equation (67), $F_m(\bar{P}_m, W_{m-1})$ increases with θ_m. What remains to prove is that $a_m = \arg\max_k L(\Theta_{m-1}, \bar{v}_k^{m-1})$ maximizes θ_m. We prove by contradiction. Let $k^* \neq a_m$, we have $L(\Theta_{m-1}, \bar{v}_{k^*}^{m-1}) < L(\Theta_{m-1}, \bar{v}_{a_m}^{m-1})$ by definition, and $U(\Theta_{m-1}, \bar{v}_{k^*}^{m-1}) \leq L(\Theta_{m-1}, \bar{v}_{a_m}^{m-1}) \leq U(\Theta_{m-1}, \bar{v}_{a_m}^{m-1})$ by characteristics. Denote θ_m by $\Psi(k)$ where k is the user selection in slot m. According to equation (57), $\Psi(k) < \Psi(a_m) \forall k \neq a_m$. Therefore, $a_m = \arg\max_k L(\Theta_{m-1}, \bar{v}_k^{m-1})$ maximizes θ_m and therefore $F_m(\bar{P}_m, W_{m-1})$.

D. Optimal Power selection

At the base case, we would like to maximize the goodput in the last slot M which is to solve

$$F_M^{(1)}(\bar{P}_M, W_{M-1}) = \max_{p_M, r_M} (1 - \epsilon) r_M. \tag{58}$$

And given W_{M-1}, θ_m can be solved by taking an inverse of the function $\phi_{a_m}(.)$ in equation (44)

$$\dot{\theta}_m = \phi^{-1} \left(\epsilon \phi(U(\Theta_{m-1}, \bar{v}_{a_m}^{m-1})) + (1 - \epsilon)\phi(L(\Theta_{m-1}, \bar{v}_{a_m}^{m-1})) \right) \tag{59}$$

As the relation of power and rate is $\theta_M = \left(\frac{N}{p_M}\right)^D 2^{\frac{DMr_M}{NT}}$, the optimal solution at the base case is

$$\begin{cases} p_M = \bar{P}_M \\ r_M = \frac{NT}{M} \log_2\left(\frac{\bar{P}_M}{N}\right) + \frac{NT}{DM} \log_2(\theta_M) \end{cases} \tag{60}$$

Therefore, r_m can be solely expressed by θ_m and p_m. Recursively develop $F_1(P_0)$, we have

$$F_1^{(1)}(P_0, W_0) = \max_{p_1, r_1} \left\{ (1 - \epsilon) r_1 + \cdots + (1 - \epsilon)^M r_M \right\} \tag{61}$$

With some mathematic manupulation, we obtain equation (68).

As we have assumed $v_m = 1$, θ_m can be computed for $m = 1$ to M. Note that $p_{m+1}, \dots, r_M$ are of the form

$$\begin{aligned} p_{m+1} &= a_1(\bar{P}_m - p_m) \\ p_{m+2} &= a_2(\bar{P}_m - p_m - p_{m+1}) \\ &\vdots \\ p_M &= a_{M-m}(1 - a_{M-m-1}) \cdots (1 - a_1)(\bar{P}_m - p_m) \end{aligned} \tag{62}$$

Therefore, the closed form of optimal power allocation is obtained. Note that the objective function in (68) is concave in

$$F_m(\bar{P}_m, W_{m-1}) = \max_{\mathbf{p}_m, \mathbf{r}_m, \mathbf{a}_m} E_{\mathbf{v}_m} E_{\hat{V}_{m+1}|\mathbf{v}_m} \left\{ Pr(c(p_m, X_{a_m}) > r_m)r_m + \sum_{i=m+1}^{M} Pr(c(p_i, X_{a_i}) > r_i)r_i \right\}. \tag{63}$$

$$F_m(\bar{P}_m, W_{m-1}) = \max_{\mathbf{p}_m, \mathbf{r}_m, \mathbf{a}_m} \left\{ Pr(c(p_m, X_{a_m} > r_m)|W_{m-1})r_m + E_{\mathbf{v}_m} \left\{ E_{\hat{V}_{m+1}|\mathbf{v}_m} \sum_{i=m+1}^{M} Pr(c(p_i, X_{a_i}) > r_i|W_{i-1})r_i \right\} \right\} \tag{64}$$

$$F_m(\bar{P}_m, W_{m-1}) = \max_{\mathbf{p}_m, \mathbf{r}_m, \mathbf{a}_m} \left\{ Pr(c(p_m, X_{a_m}) > r_m|W_{m-1})r_m + E_{\mathbf{v}_m} F_{m+1}(\bar{P}_m - p_m, W_m) \right\}. \tag{65}$$

$$F_m(\bar{P}_m, W_{m-1}) = \max_{\substack{a_m, \ldots, a_M \\ p_m, \ldots, p_M, r_m, \ldots, r_M}} \left\{ (1-\epsilon)\frac{NT}{DM} \log_2 \left(\left(\frac{p_m}{N}\right)^D \theta_m \right) + (1-\epsilon)F_{m+1}(\bar{P}_m - p_m, W_m|v_m = 1) \right\} \tag{66}$$

$$F_m(\bar{P}_m, W_{m-1}) = \max_{\substack{a_m, \ldots, a_M \\ p_m, \ldots, p_M, r_m, \ldots, r_M}} \left\{ (1-\epsilon)\frac{NT}{DM} \log_2 \left(\left(\frac{p_m}{N}\right)^D \theta_m \right) + \cdots + (1-\epsilon)^{M-m+1}\frac{NT}{DM} \log_2 \left(\left(\frac{p_M}{N}\right)^D \theta_M \right) \right\} \tag{67}$$

$$F_m^{(1)}(\bar{P}_m, W_{m-1}) = \max_{p_m, \ldots, p_M, r_m, \ldots, r_M} \left\{ \frac{NT}{M}(1-\epsilon) \left[\log_2 \left(\frac{p_m}{N}\right) + \cdots + (1-\epsilon)^{M-m} \log_2 \left(\frac{p_M}{N}\right) \right] \right. \tag{68}$$

$$\left. \frac{NT}{DM}(1-\epsilon) \left[\log_2(\theta_m) + \cdots + (1-\epsilon)^{M-m} \log_2(\theta_M) \right] \right\}$$

---

p_m. Substitute equation (62) to $F_m^{(1)}(\bar{P}_m, W_{m-1})$ in equation (68) and differentiate it and set it to zero. We obtain

$$p_m = \frac{\epsilon \bar{P}_m}{1 - (1-\epsilon)^{M-m+1}} \tag{69}$$

which is solely depending on ϵ and $\bar{P}_m$ but nothing else. The solutions obtained here is a lower bound of the original solution as the objective is solving the problem in only one direction which assumes all positive feedbacks and correspond to the all positive routes in the decision tree.

E. Shrinking User Selection Set $\mathbb{K}_m$

Before proving this lemma, we need to introduce two properties of the lower bound of channel power X_k, $L(\Theta_m, \bar{v}_k^m)$.

1) Monotonic Increasing Lower Bound of Real Channel Power:

Lemma 9: The lower bound of the channel power gains $L(\Theta_m, \bar{v}_k^m)$ increases monotonically with m.

Proof:

$$L(\Theta_m, \bar{v}_k^m) \tag{70}$$
$$= \max\{\theta_i : v_{k,i} = 1, 1 \le i \le m\}$$
$$= \begin{cases} \max\{\theta_m, \{\theta_i : v_{k,i} = 1, 1 \le i \le m-1\}\} & \text{if } v_{k,m} = 1, \\ \max\{\theta_i : v_{k,i} = 1, 1 \le i \le m-1\} & \text{if } v_{k,m} = 0. \end{cases}$$
$$= \begin{cases} \max\{\theta_m, L(\Theta_{m-1}, \bar{v}_k^{m-1})\} & \text{if } v_{k,m} = 1, \\ L(\Theta_{m-1}, \bar{v}_k^{m-1}) & \text{if } v_{k,m} = 0. \end{cases}$$
$$\ge L(\Theta_{m-1}, \bar{v}_k^{m-1})$$

∎

2) Lower Bound of Channel Power of Selected User Larger than the upper bound of channel power of the Remaining Users:

Lemma 10: Assume $\exists k \notin \mathbb{K}_{m-1}$.

$$U(\Theta_{m-1}, \bar{v}_k^{m-1}) \le L(\Theta_{m-1}, \bar{v}_{k'}^{m-1}) \ \forall k' \in \mathbb{K}_{m-1} \tag{71}$$

Proof: Assume $\exists k \notin \mathbb{K}_{m-1}$. Recall equation (41),

$$U(\Theta_{m-1}, \bar{v}_k^{m-1}) = \min\{\theta_i : v_{k,i} = 0, 1 \le i \le m-1\}$$

There exist a packet slot q, $1 \le q \le m-1$, such that $v_{k,q} = 0$ and $v_{k',q} = 1$, which can be described mathematically in equation (77).

From definition, $\theta_q \ge L(\Theta_{q-1}, v_{k'}^{q-1})$ and $L(\Theta_q, v_{k'}^q) = \max\left\{\theta_q, L(\Theta_{q-1}, v_{k'}^{q-1})|v_{k',q} = 1\right\}$. Thus, $L(\Theta_q, v_{k'}^q) = \theta_q$ if $v_{k',q} = 1$. Thus, continuing from equation (77)

$$U(\Theta_{m-1}, \bar{v}_k^{m-1}) \tag{72}$$
$$= \min\left\{L(\Theta_q, \bar{v}_{k'}^q), U(\Theta_{q-1}, \bar{v}_k^{q-1}), U(\Theta_{m-1}, \bar{v}_k^{m-1})\right\}$$
$$= L(\Theta_q, \bar{v}_{k'}^q)$$
$$\le L(\Theta_{m-1}, \bar{v}_{k'}^{m-1})$$

The last inequality is proved by lemma 9. ∎

We are going to prove this lemma by contradiction. Assume $\exists k \in \mathbb{K}_m$ and $k \notin \mathbb{K}_{m-1}$. At slot m, $\forall k' \in \mathbb{K}_{m-1}, k' \notin \mathbb{K}_m$, by lemma 9, the lower bound of channel power gain is monotonically increasing with m.

$$L(\Theta_m, \bar{v}_{k'}^m) \ge L(\Theta_{m-1}, \bar{v}_{k'}^{m-1}) \tag{73}$$

Also, by lemma 10, all users outside the user selection set have upper bound less than or equal to that of users inside the user selection set. $\forall k \notin \mathbb{K}_{m-1}, k' \in \mathbb{K}_{m-1}$

$$U(\Theta_{m-1}, \bar{v}_k^{m-1}) \le L(\Theta_{m-1}, \bar{v}_{k'}^{m-1}) \tag{74}$$

Because $k \in \mathbb{K}_m, k' \notin \mathbb{K}_m$, we have

$$L(\Theta_m, \bar{v}_k^m) > L(\Theta_m, \bar{v}_{k'}^m). \tag{75}$$

$$U(\Theta_{m-1}, \bar{v}_k^{m-1}) = \min\{\theta_q, \{\theta_i : v_{k,i} = 0, 1 \leq i \leq q-1, q+1 \leq i \leq m-1\}\} \tag{76}$$
$$= \min\left\{\theta_q, U(\Theta_{q-1}, \bar{v}_k^{q-1}), U(\Theta_{m-1}, \bar{v}_k^{m-1})\right\}$$

Thus, we have

$$L(\Theta_m, \bar{v}_k^m) \tag{77}$$
$$> L(\Theta_m, \bar{v}_{k'}^m) \quad (\forall k \in \mathbb{K}_m, k' \notin \mathbb{K}_m)$$
$$\geq L(\Theta_{m-1}, \bar{v}_{k'}^{m-1}) \quad \text{(by lemma 7)}$$
$$\geq U(\Theta_{m-1}, \bar{v}_k^{m-1}) \quad (\forall k' \in \mathbb{K}_{m-1}, k \notin \mathbb{K}_{m-1})$$

which leads to a contradiction. Thus, $\forall k \in \mathbb{K}_m, k \in \mathbb{K}_{m-1}$.

F. Channel Estimate of Selected User between Upper and Lower Bound

We are going to prove this claim by mathematical induction. In the base case, $m = 0$, before any transmission, we have initialization

$$L = 0 \tag{78}$$
$$U = \infty \tag{79}$$
$$X_k \in [L, U] \quad \forall k \in \mathbb{K}_0 \tag{80}$$

where $\mathbb{K}_0 = \{1, \ldots, K\}$.

Assume the statement is true for $m = q$. We obtain

$$X_k \in [L(\Theta_q, \bar{v}_k^q), U(\Theta_q, \bar{v}_k^q)], \ \forall k \in \mathbb{K}_q \tag{81}$$

When $m = q + 1$, before the $(q + 1)$-th transmission,

$$L(\Theta_q, \bar{v}_k^q) \leq \theta_{q+1} \leq U(\Theta_q, \bar{v}_k^q), \ \forall k \in \mathbb{K}_q \tag{82}$$

After $(q+1)$-th transmission, there are two cases, either ACK or NAK. If an ACK is received then we have

$$r_{q+1} \leq c(p_{q+1}, X_k) \tag{83}$$
$$\text{or} \quad \frac{NT}{DM} \log_2\left(\left(\frac{p_{q+1}}{N}\right)^D \theta_{q+1}\right)$$
$$\leq \frac{NT}{DM} \log_2\left(\left(\frac{p_{q+1}}{N}\right)^D X_k\right)$$
$$\text{or} \quad \theta_{q+1} \leq X_k.$$

The updates of the bounds are

$$L(\Theta_{q+1}, \bar{v}_k^{q+1}) = \max\{L(\Theta_q, \bar{v}_k^q), \theta_{q+1}\} \tag{84}$$
$$= \theta_{q+1}$$
$$\text{and} \quad U(\Theta_{q+1}, \bar{v}_k^{q+1}) = U(\Theta_q, \bar{v}_k^q). \tag{85}$$

Thus, we have $\forall k \in \mathbb{K}_q \bigcap \{k : v_{k,q+1} = 1\}$

$$L(\Theta_{q+1}, \bar{v}_k^{q+1}) \leq X_k \leq U(\Theta_{q+1}, \bar{v}_k^{q+1}). \tag{86}$$

Let $\mathbb{K}_{q+1} = \mathbb{K}_q \bigcap \{k : v_{k,q+1} = 1\}$ which completes the proof. Similarly, if NAK is received, $X_k \leq \theta_{q+1}$. The updates of the bounds are

$$L(\Theta_{q+1}, \bar{v}_k^{q+1}) = L(\Theta_q, \bar{v}_k^q) \tag{87}$$
$$U(\Theta_{q+1}, \bar{v}_k^{q+1}) = \theta_{q+1} \tag{88}$$

Thus, we have $\forall k \in \mathbb{K}_q \bigcap \{k : v_{k,q+1} = 0\}$

$$L(\Theta_{q+1}, \bar{v}_k^{q+1}) \leq X_k \leq U(\Theta_{q+1}, \bar{v}_k^{q+1}). \tag{89}$$

Let $\mathbb{K}_{q+1} = \mathbb{K}_q \bigcap \{k : v_{k,q+1} = 0\}$ which completes the proof.

G. Monotonic Decreasing Gap between Upper and Lower Bounds

The difference of the gaps at slot m and $m - 1$ is

$$w_m - w_{m-1} \tag{90}$$
$$= \{U(\Theta_m, \bar{v}_{a_m}^m) - L(\Theta_m, \bar{v}_{a_m}^m)\}$$
$$\quad - \left\{U(\Theta_{m-1}, \bar{v}_{a_{m-1}}^{m-1}) - L(\Theta_{m-1}, \bar{v}_{a_{m-1}}^{m-1})\right\}$$
$$= \begin{cases} L(\Theta_{m-1}, \bar{v}_{a_{m-1}}^{m-1}) - L(\Theta_m, \bar{v}_{a_m}^m) & \text{if } \bar{v}_{a_m}^m = 1, \\ U(\Theta_m, \bar{v}_{a_m}^m) - U(\Theta_{m-1}, \bar{v}_{a_{m-1}}^{m-1}) & \text{if } \bar{v}_{a_m}^m = 0. \end{cases}$$
$$= \begin{cases} L(\Theta_{m-1}, \bar{v}_{a_{m-1}}^{m-1}) - \theta_m & \text{if } \bar{v}_{a_m}^m = 1, \\ \theta_m - U(\Theta_{m-1}, \bar{v}_{a_{m-1}}^{m-1}) & \text{if } \bar{v}_{a_m}^m = 0. \end{cases}$$
$$\leq 0 \ \forall m$$

The last inequality is due to the fact that $L(\Theta_{m-1}, \bar{v}_{a_{m-1}}^{m-1}) \leq \theta_m \leq U(\Theta_{m-1}, \bar{v}_{a_{m-1}}^{m-1})$

H. Scheduled Rate Achieves Capacity

By lemma 5, when $m \to \infty$, the user selection set degenerates to a single user whose has the largest lower bound of the channel power gains, $\mathbb{K}_m = k$ where $L(\Theta_m, \bar{v}_k^m) > L(\Theta_m, \bar{v}_{k'}^m)$ and $k \neq k'$. Using lemma 6 and 7, we have

$$m \to \infty, \ L(\theta_m, \bar{v}_{a_m}^m) = U(\theta_m, \bar{v}_{a_m}^m) = X_{a_m} \tag{91}$$

Thus, we have

$$m \to \infty, \ \mathbb{K}_m = a_m = k, \ \text{where } X_k > X_{k'}, \ k \neq k' \tag{92}$$

Also by lemma 7, we have

$$m \to \infty, \ \theta_m = L(\theta_m, \bar{v}_{a_m}^m) = U(\theta_m, \bar{v}_{a_m}^m) = X_{a_m} \tag{93}$$

Thus, we have the scheduled rate at slot m,

$$\lim_{m \to \infty} r_m = \lim_{m \to \infty} \frac{NT}{DM} log_2\left(\left(\frac{p_m}{N}\right)^D \theta_m\right) \tag{94}$$
$$= \lim_{m \to \infty} \frac{NT}{DM} log_2\left(\left(\frac{p_m}{N}\right)^D X_k\right)$$
$$= c(p_m, X_k)$$

where user k has the largest channel power gains. The quantity $c(p_m, X_k)$ is the capacity achieved by the system with perfect CSIT.

REFERENCES

[1] A. Annamalai, L. Freiberg and V. K. Bhargava, "Analysis and Optimization of an Adaptive Go-Back-N ARQ Protocol for Time-Varying Channels",IEEE Transactions on Communications,Vol. 46, No.10, Oct 1998.

[2] A. Annamalai, V. K. Bjargava, "Analysis and Optimization of Adaptive Multicopy Transmission ARQ Protocols for Time-Varying Channels", IEEE Transactions on Communications, Vol. 46,No.10, Oct 1998.

[3] H. Minn, M. Zeng and V. K. Bhargava, "On ARQ Scheme With Adaptive Error Control", IEEE Transactions on Venhicular Technology, Vol.50, No.6, Nov 2001.

[4] J. L. Doob,*Stochastic Processes*. New York:Wiley,1953.

[5] R. Negi and J. Cioffi, "Outage Capacity with Causal Feedback", IEEE Transactions on Information Theory, Vol.48, No. 9, Sept. 2002.

[6] M. Realp, A. I. Perez-Neira, C. Mecklenbrauker, "A cross-layer approach to multi-user diversity in heterogeneous wireless systems",IEEE International Conference on Communications 2005,Volume 4, 16-20 May 2005 Page(s):2791 - 2796.

[7] Rui Wang and V. K. N. Lau, "Cross Layer Design of Downlink Multi-Antenna OFDMA Systems with Imperfect CSIT for Slow Fading Channels", IEEE Transactions on Wireless Communications, Page(s):2417 ? 2421, July 2007

[8] G. Corral-Briones, A. A. Dowhuszko, J. Hamalainen and R. Wichman, "Achievable data rates for two transmit antenna broadcast channels with WCDMA HSDPA feedback information",IEEE International Conference on Communications 2005,Vol. 4, 16-20 May 2005 Page(s):2722 - 2727.

[9] Ka Ming Ho, V. K. N. Lau, S. K. Cheng , "Closed Loop Cross-Layer Scheduling for Goodput Maximization with no CSIT",IEEE Globecom 2006.

[10] Zuleita K. M. Ho, V. K. N. Lau, S. K. Cheng, "Closed Loop Cross-Layer Scheduling For Goodput Maximization in Frequency Selective Environment with no CSIT", IEEE WCNC 2007.

[11] K. Li and X. Wang, " Cross-Layer Optimization for LDPC-Coded Multirate Multiuser Systems With QoS Constraints", IEEE Transactions on Signal Processing, Vol. 54, No. 7, July 2006.

[12] Q. Du and X. Zhang, "Cross-Layer Resource-Consumption Optimization for Mobile Multicast in Wireless Networks", IEEE International Symposium on a World of Wireless, Mobile and Multimedia 2006.

[13] A. Sali, A. Widiawan, S. Thilakawardana, R. Tafazolli and B. G. Evans,"Cross-Layer Design Approach for Multicast Scheduling over Stellite Networks", IEEE IWSSC05, Sienna, Italy, 8 Sept, 2005.

[14] Q. Liu, S. Zhou and G. B. Giannakis, "Cross-Layer Scheduling with Prescribed QoS Guarantees in Adaptive Wireless Network", IEEE Journal on Selected Areas in Communications, Vol.23, No. 5, May 2005.

[15] F. Rey, M. Lamarca and G. Vazquez, "Robust Power Allocation Algorithms for MIMO OFDM Systems with Imperfect CSI", IEEE Transactions on Signal Processing, Vol.53, No.3, March 2005.

[16] J. Tang and X. Zhang, "Space-time diversity versus feedback-based channel adaptation in cross-layer design of wireless networks", 2005 IEEE International Conference on Electro Information Technology, 22-25 May 2005.

[17] M. A. Haleem and R. Chandramouli, "Adaptive downlink scheduling and rate selection: a cross-layer design", IEEE Journal on Selected Areas in Communications, Vol. 23, Issue 6, June 2005.

[18] A. Pascual-Iserte, A. I. Perez-Neira and M. A. Lagunas, "On Power Allocation Strategies for Maximum Signal to Noise and Interference Ratio in an OFDM-MIMO System", IEEE Transactions on Wireless Communications, Col.3, No.3, May 2004.

[19] J. Tang, X. Zhang and Q. Du, "Alamouti Scheme with Joint Antenna Selection and Power Allocation over Rayleigh Fading Channels in Wireless Networks", IEEE Globecom 2005.

[20] F. Yu, V. Krishnamurthy and V. C. M. Leung, "Cross-Layer Optimal Connection Admission Control for Variable Bir Rate Multimedia Traffic in Packet Wireless CDMA Networks", IEEE Transactions on Signal Processing, Vol. 54, No. 2, Feb 2006.

[21] H. Choi and S. Park, "Analysis of Energy Efficiency for Low Power Transmission in WLAN with MIMO-OFDM and Block Ack Mechanism", IEEE ICACT 2006.

[22] J. Zhang, Z. Fang and B. Bensaou, "Adaptive Power Control for Single Channel Ad Hoc Networks", IEEE ICC 2004.

[23] W. Lilakiatsakun and A. Seneviratne, "TCP Performances over Wireless Link Deploying Delayed ACK", IEEE VTC 2003.

[24] S. H. Hwang, B. H. Kim and Y. S. Kim, "A Hybrid ARQ Scheme with Power Ramping", IEEE VTC 2001.

[25] M. A. Haleem and R. Chandramouli, "Adaptive Downlink Scheduling and Rate Selection: A Cross-Layer Design", IEEE Journal on Selected Areas in Communications, Vol. 23, No. 6, June 2006.

[26] A. K.Karmokar and Vijay K. Bhargava, "Optimal Packet Scheduling over Correlated Nakagami-m Fading Channels with Different Diversity-Combining Techniques", IEEE Globecom 2005.

[27] A. K. Karmokar and Vijay K. Bhargava, "Coding Rate Adaptation for Hybrid ARQ Systems over Time Varying Fading Channels with Partially Observble State", IEEE ICC 2005.

[28] M. L. Puterman, *Markov Decision Process- Discrete Stochastic Dynamic Programming*, John Wiley and Sons, Inc.

[29] E. Altman, *Constrained Markov Decision Processes*, Chapman & Hall/CRC, 1998.

[30] D. V. Djonin and V. Krishnamurthy, "*Q*- Learning Algorithms for Constrained Markov Decision Processes with Randomized Monotone Policies: Application to MIMO Transmission Control", IEEE Transactions on Signal Processing, Vol. 55, No. 5, May 2007.

[31] V. W. S. Wong, M. E. Lewis and V. C. M. Leung, "Stochastic Control of Path Optimization for Inter-Switch Handoffs in Wireless ATM Networks", IEEE/ACM Transactions on Networking, Vol. 9, No. 3, June 2001.

[32] C. E. Perkins and E. M. Royer. Ad-Hoc on-demand distance vector routing. Proceedings of the 2nd IEEE Workshop on Mobile computer systems and Applications, pages 90-100, 1999.

[33] Otto Rasanen. "Semi-Markov decision processes." Seminar on MDP. Nov 11, 2006 ¡http://www.cs.helsinki.fi/u/hyu/Notes/OR-smdp-presentation.pdf¿.

[34] E. A. Feinberg and A. Shwartz, *Handbook of Markov Decision Processes- Methods and Applications*, Stanford University, 2002.

[35] S. Sarkar, "Optimum scheduling and memory management in input queued switches with finite buffer space", IEEE Transactions on Information Theory, Vol. 50, Issue 12, PP3197-3220, Dec. 2004.

[36] I. Bettesh and S. Shamai, "Optimal Power and Rate Control for Minimal Average Delay: The Single-User Case", IEEE Trans. on IT, Vol 52, No. 9, Sept 2006.

[37] S. Bhardwaj, R.J. Williams and A.S. Acampora, " On the Performance of a Two-User MIMO Downlink System in Heavy Traffic", IEEE Trans. on IT, Vol. 53, No.5, May 2007.

Zuleita K. M. Ho (S'05) enrolled into Hong Kong University of Science and Technology through the Early Admission for Outstanding Secondary Six Students in 2001and graduated from the Dept of ECE, with a B.Eng (Distinction 1st Hons) and M.Phil in 2004 and 2006. She is currently a Ph.D candidate in mobile communications in EURECOM, France. Her research interests include cooperations in wireless networks, distributed barginaing, Game theory, crosslayer optimization, information theory and random matrices. Zuleita has received more than 10 academic scholarships including, The Croucher Foundation Scholarship 2007, The Hongkong Bank Foundation Overseas Scholarship Schemes 2004, which sponsors a full year study at Massachusetts Institute of Technology and The IEE Outstanding Student Award 2004.

Vincent K.N. Lau (M'97- SM'01) obtained B.Eng (Distinction 1st Hons) from the University of Hong Kong (1989-1992) and Ph.D. from Cambridge University (1995-1997). He was with HK Telecom (PCCW) as system engineer from 1992-1995 and Bell Labs - Lucent Technologies (NJ) as member of technical staff from 1997-2003. He then joined the Department of ECE, Hong Kong University of Science and Technology (HKUST) as Associate Professor. His current research interests include the robust and delay-sensitive cross-layer scheduling of MIMO/OFDM wireless systems with imperfect channel state information, cooperative and cognitive communications, dynamic spectrum access as well as stochastic approximation and Markov Decision Process. He served as the editor of IEEE Transactions on Wireless Communications, guest editor of IEEE Journal on Selected Areas in Communications (JSAC), IEEE Journal of Special Topics on Signal Processing, IEEE System Journal as well as EURASIP Journal on Wireless Communications and Networking.

Roger S.-K. Cheng (S'86-M'92) received the B.S. degree from Drexel University, Philadelphia, PA, in 1987, and the M.A. and Ph.D. degrees from Princeton University, Princeton, NJ, in 1988 and 1991, respectively, all in electrical engineering. From 1987 to 1991, he was a Research Assistant in the Department of Electrical Engineering, Princeton University, Princeton, NJ. From 1991 to 1995, he was an Assistant Professor in the Electrical and Computer Engineering Department, University of Colorado at Boulder. In June 1995, he joined the Faculty of the Hong Kong University of Science and Technology, where he is currently an Associate Professor in the Department of Electrical and Electronic Engineering. He has also held visiting positions with Qualcomm, San Diego, CA, in the summer of 1995, and with the Institute for Telecommunication Sciences, NTIA, Boulder, CO, in the summers of 1993 and 1994. His current research interests include wireless communications, OFDM, space?time processing, CDMA, digital implementation of communication systems, wireless multimedia communications, information theory, and coding. Dr. Cheng is currently an Editor for Wireless Communication for the IEEE TRANSACTIONS ON COMMUNICATIONS. He has served as Guest Editor of the special issue on Multimedia Network Radios in the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, Associate Editor of the IEEE TRANSACTION ON SIGNAL PROCESSING, and Membership Chair for of the IEEE Information Theory Society. He is the recipient of the Meitec Junior Fellowship Award from the Meitec Corporation in Japan, the George Van Ness Lothrop Fellowship from the School of Engineering and Applied Science in Princeton University, and the Research Initiation Award from the National Science Foundation